

ENTWICKLUNG UND ERPROBUNG VON AUFGABEN
ZUR ERFASSUNG ZENTRALER KOMPETENZEN IM
CHEMIEUNTERRICHT AM BEISPIEL SÄUREN UND BASEN

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

vorgelegt von
Dorothee Christiansen

Kiel
November 2007

Referent/Referentin:

Korreferent/Korreferentin:

Tag der mündlichen Prüfung:

Zum Druck genehmigt:

Kiel, den

Der Dekan

ERKLÄRUNG

Hiermit erkläre ich, dass die vorliegende Dissertation, abgesehen von der Beratung durch meine Betreuerin, Prof. Dr. Ilka Parchmann, nach Inhalt und Form mein eigenes Werk ist. Bei der Anfertigung dieser Arbeit habe ich keine anderen als die angegebenen Quellen und Hilfsmittel verwendet.

Darüber hinaus erkläre ich, dass ich noch keinen Promotionsversuch unternommen habe und dass insbesondere diese Arbeit weder vollständig noch in Teilen an anderer Stelle im Rahmen eines Prüfungsverfahrens vorgelegt, veröffentlicht oder zur Veröffentlichung eingereicht wurde.

Kiel, den 14. November 2007

ZUSAMMENFASSUNG

Die Umsetzung innovativer Konzepte für den naturwissenschaftlichen Unterricht einerseits und die Einführung kompetenzorientierter Bildungsstandards andererseits machen eine Auseinandersetzung mit der Konstruktion von Aufgaben notwendig, die den Anforderungen eines veränderten Unterrichts und der Kompetenzerfassung genügen. Auch die Realisierung der Unterrichtskonzeption *Chemie im Kontext*, die seit 2002 im Rahmen eines Projektes zur Untersuchung von Einflussfaktoren auf den Implementationserfolg in Schulen zahlreicher Bundesländer weiterentwickelt und erprobt wird, wirft die Frage auf, inwieweit sich über Fachwissen hinausgehende Ziele des innovativen Chemieunterrichts erfassen und beurteilen lassen.

In der Arbeit wurde als zentrale Fragestellung verfolgt, ob sich anhand eines theoretisch ausgearbeiteten Modells zum naturwissenschaftlichen Denken und Arbeiten in einem Unterricht nach *Chemie im Kontext* Testaufgaben konstruieren lassen, die statistischen Qualitätsansprüchen genügen und differenzierende Aussagen über Lernstände von Schülerinnen und Schülern hinsichtlich verschiedener Kompetenzen des Konstrukts „Scientific Literacy“ ermöglichen.

Konkretes Ziel der Studie war die Entwicklung von kompetenzorientierten Testaufgaben zur Lernstandserfassung zum exemplarisch ausgewählten Thema „Säuren, Basen und Neutralisation“. Die Arbeit gliederte sich in eine Pilot- und eine Hauptuntersuchung. In der Pilotphase nahmen 16 Lernende an einer mündlichen Gruppenbearbeitung und 120 Probanden an einer Fragebogenerhebung teil. In der Hauptuntersuchung bearbeiteten 359 Lernende die Testaufgaben; die 17 unterrichtenden Lehrkräfte wurden darüber hinaus u. a. zu dem vorangegangenen Unterricht und zu ihrer Einschätzung der Schülerleistungen befragt.

Ergebnis der Pilotstudie war ein überarbeitetes Instrument für die Hauptuntersuchung. Die Hauptuntersuchung zeigte zunächst, dass das Testinstrument nicht die theoretisch zu Grunde gelegten Kompetenzbereiche bestätigte, sondern dass der Test zum Themenbereich Säuren und Basen insgesamt ein eindimensionales Instrument mit guten statistischen Kennwerten darstellte.

Mit dem Test konnten u. a. Aussagen zu den Leistungen verschiedener Probandengruppen in Bezug auf eine Gesamtdimension des naturwissenschaftlichen Denkens und Arbeitens getroffen werden. Insgesamt wurde festgestellt, dass sich Leistungen einzelner Klassen signifikant unterscheiden, es aber keinen signifikanten Leistungsunterschied gibt zwischen der Gruppe, die nach *Chemie im Kontext* unterrichtet wurde, und einer Gruppe, die keine Konzeptionserfahrung hat. Zwischen der Lehrereinschätzung und dem tatsächlichen Lösungsverhalten der Lernenden zeigten sich z. T. deutliche Abweichungen. Für unterrichtete Inhalte überschätzten die Lehrkräfte die Leistungen ihrer Schülerinnen und Schüler, während sie die Leistungen ihrer Lerngruppe bei eher „unüblichen“ Aufgaben (z. B. Bewertung von Werbung) unterschätzten.

ABSTRACT

Innovative concepts for science teaching and the implementation of national educational standards make it necessary to work on the construction of tasks, which have to regard both, a different kind of teaching sciences and setting priority on competencies instead of factual knowledge. In realising the innovative instructional design *Chemie im Kontext*, in which about 200 teachers are developing and adapting teaching units since 2002, the supporting and hindering factors for a successful implementation have been examined. Within this project the question of how the aims of the innovative instructional design can be measured and evaluated has also risen.

The aim of the presented thesis was to develop and to analyse tasks, which offer insight into the students' competencies in various areas of scientific literacy, dealing with one particular chemical topic in different contexts. To design the items, the German National Standards and a definition of scientific literacy were used as well as the system of cognitive competencies from the additional German PISA study. The exemplary chemical topic was "acids and bases". It was explored in two different daily-life contexts (heartburn, dental care) and as "pure" school chemistry. All tasks have been formulated in a way that they can be used in test instruments as well as in teaching units, both regarding the affordances of context-based learning and of "traditional chemistry".

The described study was divided into a pilot and a main phase. The goal of the pilot study was to test and revise the tasks, used in the final test. The revised version of the test was used as the central research instrument of the main study, in which 17 classes ($n = 359$, average age 16) participated. In addition to that, the teachers were asked to estimate their students' results for all questions and to evaluate the emphasis they had put on each specific kind of task in their unit on acid and bases.

The results showed large and significant variations between classes, whereas no significant differences could be found between classes following *Chemie im Kontext* and traditional classes. Regarding teachers' estimations, several mismatches have been found. For aspects taught in class, teachers often overestimated their students' results, while they expected less good results for "unusual" tasks, such as judgement of advertisements.

INHALTSVERZEICHNIS

1	Einleitung	1
THEORETISCHER HINTERGRUND		
2	Grundlagen der Konzeption <i>Chemie im Kontext</i>	5
2.1	<i>Chemie im Kontext</i> – Vorstellung einer kontextbasierten, ganzheitlichen Konzeption für den Chemieunterricht	5
2.1.1	Situiertes Lernen	8
2.1.2	Selbstgesteuertes Lernen	10
2.1.3	Naturwissenschaftliche Grundbildung	12
2.1.3.1	Kompetenzen und Kompetenzerfassung	12
2.1.3.2	Naturwissenschaftliche Grundbildung nach PISA	13
2.1.3.3	Nature of Science	15
2.1.3.4	Bildungsstandards im Fach Chemie	17
2.1.4	Zusammenfassung	18
2.2	Prozess der Erkenntnisgewinnung in einem Unterricht nach <i>Chemie im Kontext</i>	19
2.3	Wege in die Schulpraxis: Implementation der Konzeption <i>Chemie im Kontext</i>	24
2.3.1	Symbiotische Implementationsstrategie	24
2.3.2	Schulsets – Zusammensetzung und Aufgaben	26
2.4	Erfahrungen bei der Umsetzung der Konzeption in der Praxis	27
2.5	Neuer Unterricht – Alte Bewertungsmaßstäbe?!	30
3	Aufgaben – Gestaltung und Einsatz im naturwissenschaftlichen Unterricht	31
3.1	Aufgaben und Aufgabenkultur – Begriffsfestlegungen	32
3.2	Einsatzgebiete von Aufgaben im Unterricht	33
3.2.1	Gestaltungsmittel und Diagnosewerkzeuge im Unterricht	33
3.2.2	Bewertungsinstrument für Lernende	34
3.3	Ansätze zur Klassifizierung von Aufgaben zur Analyse und Konstruktion	35
3.3.1	Allgemeine Forderungen – Aufgabeneinsatz, Kontextorientierung und Öffnung von Aufgaben	36

3.3.2	„Gute“ Aufgaben	37
3.3.3	Systematisierung der verschiedenen Ansätze	39
3.3.3.1	Zuordnung einer Aufgabe nach Inhalt, Lehrplan oder Richtlinien	40
3.3.3.2	Zuordnung einer Aufgabe zu bestimmten Unterrichtsphasen	40
3.3.3.3	Einordnung einer Aufgabe nach Art der möglichen Lösungswege	41
3.3.3.4	Klassifizierung einer Aufgabe nach ihrem Antwortformat	42
3.3.3.5	Klassifizierung einer Aufgabe nach dem Grad ihrer Offenheit	43
3.3.3.6	Zuordnung einer Aufgabe nach dem erforderlichen Experimentierverhalten	43
3.3.3.7	Klassifizierung einer Aufgabe entsprechend ihren Anforderungsmerkmalen	44
3.3.3.8	Zuordnung einer Aufgabe zu vorgegebenen Kompetenzstufen oder Kompetenzbereichen	47
3.3.3.9	Sonderfall: Beispielaufgaben	49
3.3.3.10	Zusammenfassung	50
3.4	Was ist eine „ <i>Chemie-im-Kontext</i> -Aufgabe“? Versuch einer Definition	51
3.5	Kompetenzerfassung – Hinweise zur Testkonstruktion und Aufgabenentwicklung	52
4	Entwicklung von Kategorien zur Analyse und Konstruktion von Aufgaben, die den Anforderungen eines Unterrichts nach <i>Chemie im Kontext</i> gerecht werden	56
4.1	Inhaltliche Kriterien zur Konstruktion und Analyse von Aufgaben	56
4.1.1	Erste Dimension: Naturwissenschaftliches Denken und Arbeiten in <i>Chemie im Kontext</i>	57
4.1.2	Zweite Dimension: Kognitive Aufgabenanforderungen im Bereich der Naturwissenschaften	58
4.2	Zusammenführung zu einem Modell zur Beschreibung von Aufgaben	58
4.3	Testaufgaben der Hauptuntersuchung	62
4.3.1	Inhalt, Lehrplan und Richtlinien	63
4.3.2	Unterrichtsphasen	64
4.3.3	Art der möglichen Lösungswege	64
4.3.4	Antwortformate	65
4.3.5	Grad der Offenheit	66
4.3.6	Erforderliches Experimentierverhalten	66

4.3.7	Anforderungsmerkmale	66
4.3.8	Kompetenzbereiche, Kompetenzstufen	67
4.3.9	Überblick über die Kategorisierung aller Testaufgaben	67

EMPIRISCHE UNTERSUCHUNGEN

5	Pilotstudie: Entwicklung und Erprobung von Testaufgaben zur Erfassung zentraler Kompetenzen im Chemieunterricht	70
5.1	Ziele der Untersuchung	71
5.2	Fragestellungen	71
5.3	Methode	73
5.3.1	Stichprobenbeschreibung	74
5.3.2	Erhebungsinstrumente der Pilotstudie	75
5.3.3	Auswertung der Testitems	78
5.4	Ergebnisse	82
5.4.1	Phase 1 der Pilotuntersuchung: Videoaufzeichnung der Gruppenbearbeitung	82
5.4.2	Zwischendiskussion	83
5.4.3	Phase 2 der Pilotuntersuchung: Fragebogeneinsatz	83
5.5	Diskussion	96
5.6	Zusammenfassung	104
6	Hauptuntersuchung: Einsatz von Testaufgaben zur Erfassung zentraler Kompetenzen im Chemieunterricht	105
6.1	Ziele der Untersuchung	105
6.2	Fragestellungen	105
6.2.1	Fragestellung 1: Statistische Prüfung des Testinstruments	106
6.2.2	Fragestellung 2: Auswertung der Leistungen der Schülerinnen und Schüler	106
6.2.3	Fragestellung 3: Einzelbetrachtung ausgewählter Aufgaben	107
6.3	Methode	108
6.3.1	Beschreibung der Stichprobe	108
6.3.2	Erhebungsinstrumente der Hauptstudie	110
6.4	Ergebnisse	113
6.4.1	Ergebnisse zur Auswertung des Lehrerfragebogens	113

6.4.2	Zwischendiskussion 1	115
6.4.3	Ergebnisse der Auswertung des Schülerfragebogens	116
6.4.3.1	Statistische Prüfung des Testinstruments (Fragestellung 1)	116
6.4.3.2	Zwischendiskussion 2	122
6.4.3.3	Leistungen der Schülerinnen und Schüler (Fragestellung 2)	124
6.4.3.4	Zwischendiskussion 3	139
6.4.3.5	Einzelbetrachtung ausgewählter Aufgaben (Fragestellung 3)	142
6.4.2.6	Zwischendiskussion 4	151
6.4.3.7	Diskussion	154
7	Zusammenfassung und Ausblick	157
LITERATUR		162
ANHANG		169
A	Fragebogen der Pilotstudie	169
B	Erwartungshorizont (Pilotstudie)	178
C	Klassifizierung der Testaufgaben (Pilotstudie)	179
D	Ergänzungen zum Ergebnisteil (Pilotstudie)	181
E	Schülerfragebogen der Hauptuntersuchung	187
F	Erwartungshorizont (Hauptstudie)	196
G	Kurzbeschreibung der Testitems (Hauptstudie)	198
H	Lehrerfragebogen der Hauptuntersuchung	199
I	Ergänzungen zum Ergebnisteil (Hauptstudie)	210
K	CD-ROM: Unterrichtsmaterialien	213

1 Einleitung

Der Themenbereich Aufgaben und Kompetenzerfassung nimmt einen zentralen Platz in der aktuellen bildungspolitischen und fachdidaktischen Diskussion ein. Nicht zuletzt durch die Entscheidung zur Einführung nationaler Bildungsstandards und in einigen Bundesländern die Ein- oder erstmalige Durchführung eines Zentralabiturs ist das Formulieren von Aufgaben zur Erfassung von Fachwissen und Kompetenzen ins Blickfeld gerückt. Ausgangspunkt waren dafür u. a. die Ergebnisse der deutschen Schülerinnen und Schüler bei internationalen, vergleichenden Leistungsstudien wie TIMSS und PISA, die die Diskussion in die Wege leiteten. Es ist bekannt, dass die Leistungen deutscher Schülerinnen und Schüler in den Naturwissenschaften im Vergleich mit anderen Ländern bestenfalls im Mittelfeld anzusiedeln sind und dass darüber hinaus deutliche Unterschiede im Leistungsstand der Schülerinnen und Schüler in den verschiedenen Bundesländern festzustellen sind.

Von den Problemen des naturwissenschaftlichen Unterrichts an deutschen Schulen ist vielfach berichtet worden (u. a. Parchmann, Demuth, Ralle, Paschmann & Huntemann 2001, BLK 1997, Duit, Fischer & Müller 2002). Darüber hinaus muss der naturwissenschaftliche Unterricht die Lernenden auf eine technologisierte Welt vorbereiten, die sich ohne grundlegende Kenntnisse in den Naturwissenschaften immer weniger begreifen lässt. In der heutigen Wissensgesellschaft geht es nicht mehr so sehr darum, Informationen zu wissen, sondern sich in einem Überfluss von Informationen zurechtzufinden und die Vielzahl von Informationen und Informationsquellen auf der Basis eigener Kenntnisse und Fähigkeiten zu beurteilen. Damit wird die naturwissenschaftliche Grundbildung ein wichtiger Teil der Allgemeinbildung. Die Vermittlung von Wissen ist demnach nur noch ein Teil der Anforderungen, denen naturwissenschaftlicher Unterricht gerecht werden muss.

Verschiedene Ansätze zur Veränderung des naturwissenschaftlichen Unterrichts setzen an diesen Erkenntnissen an. Die Unterrichtskonzeption *Chemie im Kontext* stellt für den Chemieunterricht eine Möglichkeit dar, den Problemen zu begegnen, indem in einem neuen konzeptionellen Ansatz die Verbindung von situiertem Lernen mit dem Aufbau von systematisch angelegtem Wissen hergestellt werden soll (Parchmann et al. 2001). Diese Konzeption fußt auf drei grundlegenden Prinzipien: das erste Prinzip der Ausrichtung des gesamten Unterrichts an Kontexten ergibt sich u. a. aus den Theorien zum situierten Lernen und der Interessens- und Motivationsforschung, das zweite Prinzip des sukzessiven Aufbaus von grundlegenden Basiskonzepten der Chemie trägt u. a. der Forderung nach kumulativem Lernen Rechnung. Das dritte Prinzip der Vielfalt in der Unterrichtsgestaltung berücksichtigt das unterschiedliche Vorwissen der Lernen-

den und ihre individuellen Lernstrategien und legt damit u. a. den Grundstein für die Fähigkeit zum selbständigen und lebenslangen Lernen.

Der zweite Themenkomplex, der in engem Zusammenhang mit der Veränderung von Unterricht steht, ist der Bereich der Aufgaben und Aufgabenkultur. Aufgaben allein können Unterricht nicht verändern, aber eine Veränderung von Unterricht schließt ein Nachdenken über Aufgaben mit ein (Duit et al. 2002). Für Lehrkräfte sind Aufgaben im Unterricht sehr zentral, da über sie der Unterricht gestaltet wird. Für eine Analyse und Konstruktion von geeigneten Aufgaben in einem veränderten Chemieunterricht bedarf es in erster Linie transparenter Kriterien, anhand derer sich Aufgaben für einen Unterricht nach *Chemie im Kontext* analysieren und konstruieren lassen, wobei diese Kriterien aus den Zielstellungen der Konzeption abgeleitet werden müssen.

In diesem Zusammenhang wird ebenfalls deutlich, dass das Fachwissen nur noch *ein* Ziel von Unterricht sein kann, neben dem ganz maßgeblich der Aufbau und die Förderung von Kompetenzen stehen, die für das naturwissenschaftlichen Denken und Arbeiten von zentraler Bedeutung sind. In einem Unterricht, in dem nicht mehr nur das, was gelernt wird (Fachinhalte), sondern vor allem auch wie in Chemie und Naturwissenschaften allgemein gelernt und Erkenntnisse gewonnen werden, in den Mittelpunkt rückt, gilt es Wege zu finden und Instrumente zu entwickeln, die der Erfassung beider Seiten, Inhalten und Kompetenzen, gerecht werden.

Ein Ziel der vorliegenden Arbeit ist es daher, ein Modell der Erkenntnisgewinnung aus den Prinzipien und Zielvorstellungen der Konzeption für einen Unterricht nach *Chemie im Kontext* zu entwickeln, um anhand dieser Zielformulierungen konkrete Kriterien ableiten zu können, mit denen Aufgaben auf ihre Passung zu einem kontextbasierten Chemieunterricht beurteilt und konstruiert werden können.

Um den Blick auf das Themengebiet Aufgaben und Aufgabenkultur im Chemieunterricht zu vervollständigen, werden verschiedene Ansätze, die zur Analyse und Konstruktion von Aufgaben für den naturwissenschaftlichen Unterricht in der fachwissenschaftlichen Literatur existieren, zusammengestellt und auf ihre Brauchbarkeit speziell im Chemieunterricht hin bewertet.

Die vorliegende Arbeit greift damit zentrale Problemstellungen auf, die sich in der praktischen Arbeit und Umsetzung der Konzeption im Unterricht ergeben und ist in ein von Bund und Ländern gefördertes Projekt zur Implementation von *Chemie im Kontext* in den Chemieunterricht eingebunden.

Im Rahmen der Einführung von zentralen Bildungsstandards und der damit einhergehenden Schwerpunktverschiebung von dem Abprüfen von Fachwissen auf das Erfassen von Kompetenzen stellt sich auch im Zusammenhang mit der konzeptionellen und unterrichtlichen Arbeit mit

Chemie im Kontext die Frage nach geeigneten Aufgabenstellungen, die eine Kompetenzentwicklung im Bereich des naturwissenschaftlichen Arbeitens fördern.

Anhand des entwickelten Modells zur Erkenntnisgewinnung werden daher exemplarisch Aufgaben entwickelt, die anschließend daraufhin empirisch untersucht werden, ob und inwieweit sie geeignet sind, im Sinne einer Lernstandserhebung Aussagen über den Kompetenz- und Wissensstand hinsichtlich eines bestimmten chemischen Basiskonzept

Dabei wird einerseits zu klären sein, ob die Testaufgaben statistischen Qualitätskriterien genügen und ob es möglich ist, anhand des Fragebogens verschiedene Facetten des naturwissenschaftlichen Arbeitens zu identifizieren. Andererseits soll die durchgeführte Untersuchung zeigen, ob die Aufgaben dazu geeignet sind, Schülerleistungen zu differenzieren.

Aufgrund der Arbeit sollen abschließend Hinweise gegeben werden, die bei der Formulierung von Aufgaben zur Überprüfung von Kompetenzen berücksichtigt werden müssen. Darüber hinaus stellt das Modell zur Erkenntnisgewinnung eine Hilfe zur Entwicklung und Durchführung eines kontextbasierten Unterrichts dar, der sich an den Prinzipien Aufbau von zentralen Basiskonzepten und Unterrichtsgestaltung mit dem Ziel der Förderung des selbständigen Lernens von Schülerinnen und Schülern orientiert.

Die vorliegende Arbeit gliedert sich in einen theoretischen und einen empirischen Teil. Im ersten Teil der Arbeit wird das Projekt *Chemie im Kontext* in seinen Grundzügen genauer vorgestellt, dabei wird auch darüber Auskunft gegeben, wie die theoretischen Überlegungen des Ansatzes in die Praxis getragen und dort umgesetzt werden (Kapitel 2). Anhand verschiedener Ansätze zur Definition von naturwissenschaftlicher Grundbildung werden die Ziele der Konzeption inhaltlich gefüllt und konkretisiert.

Im Anschluss an die Vorstellung der Konzeption wird das Themengebiet Aufgaben und Aufgabenkultur genauer beleuchtet (Kapitel 3). Ziel ist es, verschiedene Möglichkeiten zur Analyse und Konstruktion von Aufgaben aufzuzeigen und zu systematisieren, die in der Fachdidaktik vorgeschlagen werden.

Im letzten Abschnitt des theoretischen Teils werden die Ziele der Konzeption und die Kriterien zu Aufgaben und ihrer Gestaltung zu einem Modell zusammengeführt und vorgestellt (Kapitel 4). In einem nächsten Schritt wird die Funktionalität dieses Modells überprüft, indem beispielhaft die Analyse einiger Aufgaben, die aus *Chemie-im-Kontext*-Unterrichtsmaterialien stammen, erprobt wird.

Abschließend wird eine Kategorisierung der im Rahmen der Arbeit entwickelten Testaufgaben vorgenommen, wobei der Entwicklung der Aufgaben ein ständiger Wechsel zwischen der Kons-

truktions- und der Analysearbeit zu Grunde lag, um die Aufgaben dahingehend zu optimieren, möglichst viele Facetten des naturwissenschaftlichen Arbeitens abzudecken.

Der empirische Teil der vorliegenden Arbeit gliedert sich in eine Pilot- und eine Hauptstudie:

Kapitel 5: Ziel der Pilotstudie ist es, das Aufgabenverständnis der Schülerinnen und Schüler für die entwickelten Testaufgaben zu überprüfen. Darüber hinaus werden anhand statistischer Gütekriterien Aussagen zur Qualität des Tests als Messinstrument getroffen. In diesem Zusammenhang wird exemplarisch überprüft, ob sich die Testaufgaben für eine Differenzierung von Schülerleistungen eignen. Auf der Basis der gewonnenen Erkenntnisse aus der Pilotphase wird der Fragebogen überarbeitet und die vorgenommenen Änderungen dargelegt.

Kapitel 6: Ziel der Hauptstudie ist neben der Überprüfung der Aussagekraft der Testaufgaben hinsichtlich der verschiedenen Kenntnisse und Kompetenzen im Bereich „Säuren und Basen“ die detaillierte Analyse von Schülerleistungen und damit zusammenhängend eine inhaltliche Auswertung exemplarischer Schülerantworten.

Im abschließenden Kapitel 7 werden in der Gesamtdiskussion die wichtigsten Ergebnisse zusammengefasst, ihre praktische Bedeutung für die unterrichtliche Arbeit mit *Chemie im Kontext* herausgestellt und weiterführende Forschungsfragen abgeleitet.

2 Die Unterrichtskonzeption *Chemie im Kontext*

Im Folgenden wird die Konzeption *Chemie im Kontext* vorgestellt, deren theoretische Inhalte und Ziele die Grundlage der Entwicklung eines Kompetenz- und Wissenstests im Rahmen dieser Arbeit bilden. Beginnend mit einer kurzen Darstellung der grundlegenden Prinzipien der Konzeption werden im Anschluss daran die Ziele aufgezeigt, die ein an der Konzeption orientierter Chemieunterricht verfolgt. Aus dem übergeordneten Ziel, die Lernenden „scientific literate“ aus dem Chemieunterricht zu entlassen, werden konkretere Kompetenzen abgeleitet, die für das Betreiben von Chemie im lebensweltlichen Kontext zentral sind.

Neben dem inhaltlichen Ziel des Aufbaus einer naturwissenschaftlichen Grundbildung wird auf das methodische Ziel einer Befähigung zum selbstgesteuerten Lernen eingegangen. Ausgehend von der Definition naturwissenschaftlicher Grundbildung müssen für den Unterricht, aber auch für die Entwicklung von Testinstrumenten konkretere Fähigkeitsbereiche mit den ihnen zugeordneten Fähigkeiten hergeleitet werden, was aber in diesem Abschnitt zunächst nur kurz aufgegriffen wird, um dann in einem eigenen Kapitel vertieft beschrieben zu werden (vgl. Kap. 4).

Nach den theoretischen Grundlagen und Zielen der Konzeption wird das Gesamtprojekt vorgestellt, mit dem die Implementation dieses Unterrichtskonzepts in die Schulpraxis realisiert wird und das den praktischen Rahmen für die vorliegende Arbeit liefert. In diesem Abschnitt wird ausgeführt, wie bei theoretischen Konzeptionen für den Unterricht der Weg aus der Wissenschaft und Forschung in die Schulpraxis bereits mitgedacht werden muss (Gräsel & Parchmann 2004) und der Ansatz der symbiotischen Implementation konzipiert ist, der bei der Umsetzung der Konzeption *Chemie im Kontext* in der Praxis verwirklicht wird (Gräsel et al. 2004, Fey, Gräsel, Puhl & Parchmann 2004).

Abschließend zeigt dieses Kapitel zentrale Fragestellungen und Problembereiche auf, die sich bei der Realisierung der Konzeption in der Praxis des Chemieunterrichts gezeigt haben, und stellt damit zugleich die Überleitung zu dem zweiten theoretischen Hintergrund der Entwicklung eines Kompetenztests dar: Aufgaben und Aufgabengestaltung im naturwissenschaftlichen Unterricht.

2.1 Theoretische Grundlagen einer kontextbasierten, ganzheitlichen Konzeption für den Chemieunterricht

Die Grundlagen der Konzeption *Chemie im Kontext* werden im Rahmen dieser Arbeit nur zusammengefasst, da sie in der Literatur bereits auf vielfältige Weise dargestellt sind und daher an

dieser Stelle auf weiterführende Veröffentlichungen zur Konzeption verwiesen werden kann (zu den konzeptionellen Grundlagen: Parchmann et al. 2001, Bündler, Demuth & Parchmann 2003; zu Unterrichtsbeispielen für die Sekundarstufen I und II: Huntemann, Vennemann & Parchmann 2000, Schmidt, Rebentisch & Parchmann 2003, Christiansen, Huntemann, Schmidt & Parchmann 2003; weiterführende Literatur zu speziellen Aspekten der Konzeption wie *Schülvorstellungen* Parchmann & Schmidt 2003, *Aufgabenkultur* Nentwig, Christiansen & Steinhoff 2004 und *Kompetenzerwerb* Prenzel & Parchmann 2003, Parchmann & Menthe 2004). Die nachfolgende Vorstellung der Konzeption legt den Schwerpunkt auf die übergeordnete Fragestellung der vorliegenden Arbeit nach Vermittlung und Erfassung von Kompetenzen in einem kontextbasierten Chemieunterricht.

Chemie im Kontext stellt idealerweise einen Unterrichtsgang für den gesamten Chemieunterricht vom Anfangsunterricht bis zum Abitur dar. Der Konzeption *Chemie im Kontext* basiert auf drei leitenden Prinzipien.

Kontexte. Die Kontexte liefern die thematische Grundlage für den konkreten Unterricht. Chemische Fachinhalte werden nicht um ihrer selbst Willen Gegenstand des Unterrichts, sondern stellen notwendige Bausteine auf dem Weg der Erarbeitung eines Kontextthemas dar. Kontexte sind komplexe Themengebiete, die aus der Lebenswelt der Lernenden stammen, persönliche und/oder gesellschaftliche Relevanz besitzen und eine alltäglichen Fragestellungen immanente Komplexität aufweisen (zu weiteren Ausführungen zum Kontextbegriff z. B. Parchmann et al. 2001, Nentwig, Parchmann, Demuth, Gräsel & Ralle 2002).

Basiskonzepte. Während ein bestimmter Kontext den „roten Faden“ für eine einzelne Einheit darstellt und über die gesamte Phase der Bearbeitung hinweg ständig präsent sein muss, sind die Basiskonzepte kontext-übergreifende Strukturierungselemente der Wissenschaftsdisziplin der Chemie. Sie treten damit als Ordnungsschema an die Stelle der Fachsystematik und dienen auf diese Weise als „roter Faden“ für den gesamten Chemieunterricht. Ein Unterricht nach *Chemie im Kontext* zielt auf einen sukzessiven, systematischen Aufbau weniger, grundlegender Konzepte der Chemie (Bündler et al. 2003, MNU 2000, Demuth, Ralle & Parchmann 2005, Bündler, Demuth, Freienberg, Klüter, Parchmann & Ralle 2006). Durch die Situierung des Lernens und die Kontextbasierung des Chemieunterrichts einerseits sowie die Abstraktion der erarbeiteten Inhalte zu wenigen, strukturierenden Konzepten andererseits soll der Aufbau eines anwendbaren und anschlussfähigen Wissens realisiert werden.

Unterrichtsgestaltung. Die dritte theoretische Grundlage der Konzeption umfasst die Forderung nach Vielfalt in der Unterrichtsgestaltung, d. h. eine möglichst große Auswahl in den eingesetzten Unterrichtsmethoden, den verwendeten Medien und der Lernenden- und Lehrendenaktivitäten.

ten soll umgesetzt werden. Auf diese Weise werden verschiedene Vorerfahrungen, das Vorwissen und unterschiedliche Lernstrategien der Schülerinnen und Schüler berücksichtigt. Insgesamt wird das Ziel verfolgt, auf Seiten der Schülerinnen und Schüler selbstgesteuertes und selbstorganisiertes Lernen zu fördern, um sie als eigenverantwortliche, selbständige Lerner aus der Schule zu entlassen.

In der praktischen Umsetzung der Konzeption werden Kontexte in vier aufeinander folgenden Phasen erarbeitet, die für alle Einheiten konstitutiv sind: Begegnungsphase, Phase der Planung der weiteren Erarbeitung (Planungsphase), Phase der eigentlichen Erarbeitung (Erarbeitungsphase) und abschließend die Vertiefungs- und Vernetzungsphase (Parchmann et al. 2001, vgl. auch MNU 2000).

In allen Phasen liegt der Fokus auf der Förderung der Eigentätigkeit und Selbständigkeit der Lernenden mit dem übergeordneten Ziel, die Verantwortung für das eigene Lernen zu stärken. Abbildung 1 zeigt, welche Zielsetzung in Bezug auf die Kontexterarbeitung in den einzelnen Phasen verfolgt wird und welche Methoden und Unterrichtsformen dabei beispielhaft zur Anwendung kommen können.

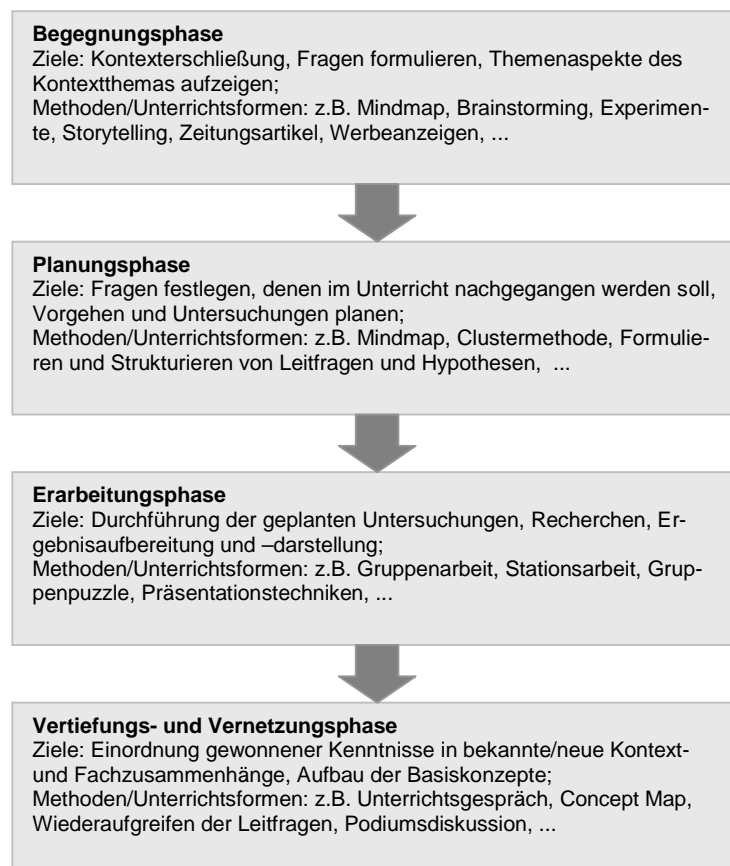


Abb. 1: Gestaltung eines Unterrichts nach *Chemie im Kontext*: Ziele der vier Phasen.

2.1.1 Situiertes Lernen

Die drei genannten Prinzipien Kontextorientierung, Aufbau von Basiskonzepten und Vielfalt in der Unterrichtsgestaltung leiten sich aus Erkenntnissen der Lehr-Lern-Forschung ab. Lernen wird nicht als eine Übertragung von Inhalten, sondern entsprechend einem konstruktivistischen Grundverständnis als ein aktiver und individueller Konstruktionsprozess des einzelnen Lernenden verstanden (Dubs 1995, Duit 1995, Terhart 1999).

Nach Ansätzen des situierten Lernens bleibt erworbenes Wissen zunächst immer an den Kontext gebunden, in dem es gelernt wird, kann aber durch häufiges Anwenden der Wissensbestände in verschiedenen Bereichen vom eigentlichen Kontext gelöst werden (für eine ausführliche Darstellung zum konzeptionellen Hintergrund des Wissenserwerbs in einem Unterricht nach *Chemie im Kontext* vgl. Steinhoff 2004). Für einen Unterricht nach *Chemie im Kontext* bedeutet dies, dass Fachinhalte, die in einem bestimmten Kontext erarbeitet werden, diesem Kontext zunächst auch verhaftet bleiben, dass sie aber durch wiederholtes Thematisieren in neuen, anderen Zusammenhängen schließlich unabhängig vom ursprünglichen Lernkontext werden.

Ein Unterricht nach *Chemie im Kontext* ist folglich durch eine Aufeinanderfolge verschiedener Kontexte gekennzeichnet, in denen jeweils bereits erarbeitete Inhalte erneut aufgegriffen und neue Inhalte erarbeitet werden. Diese z. T. von Schülerinnen und Schülern als „unübersichtlich“ empfundene Vielfalt von Fachinhalten wird in einem Strukturierungsschema zusammengeführt, indem die Inhalte den fünf Basiskonzepten zugeordnet und auf diese Weise in ein strukturiertes Wissensgebäude eingefügt werden (Bünder et al. 2003).

Das im Unterricht erworbene Wissen erlangt für die Lernenden Bedeutung, da es in Kontexte eingebettet ist, die sich für die Lernenden entweder durch persönliche Relevanz auszeichnen oder den Lernenden eine Partizipation an gesellschaftlich bedeutenden Diskussionen und Prozessen ermöglicht. Den Schülerinnen und Schülern kommt in einem Unterricht nach *Chemie im Kontext* eine sehr viel aktivere Rolle zu, sie lernen, Verantwortung im Lernprozess zu übernehmen, und haben mehr Einfluss auf die Gestaltung und Durchführung des Unterrichts, indem sie explizit an der Planung beteiligt sind. Nicht nur durch die Anbindung des Unterrichts an Themen aus der Lebenswelt der Schülerinnen und Schüler, sondern auch durch die verstärkte Schülerbeteiligung am Planungs- und Durchführungsprozess von Unterricht können Motivation und Interesse der Lernenden am Unterrichtsfach und am Unterrichtsgegenstand gefördert werden (Deci & Ryan 1993, Krapp 1993).

Die Forderungen und Hinweise für Lernumgebungen, die nach konstruktivistischen Maßstäben gestaltet sind, haben Gerstenmaier & Mandl (1995) aus verschiedenen instruktionalen Ansätzen (dem Anchored Instruction-Ansatz (The Cognition and Technology Group at Vanderbilt 1996),

dem Cognitiv Flexibility- (Spiro & Jehng 1990) und dem Cognitive Apprenticeship-Ansatz (Collins, Brown & Newman 1989)) hergeleitet (Gerstenmaier et al. 1995, ebenfalls dazu Mandl, Gruber & Renkl 1995, Dubs 1995).

Diese Forderungen kommen auch in den theoretischen Grundlagen von *Chemie im Kontext* zum Tragen, wie z. B. in der Authentizität von Lernumgebungen durch realistische Fragestellungen und Situationen sowie in einem subjektiv von den Lernenden wahrnehmbaren Handlungsspielraum im Unterricht. Zudem werden Themen aus verschiedenen Blickrichtungen und Standpunkten heraus bearbeitet und präsentiert, um auch auf diese Weise eine Flexibilisierung des Denkens der Schülerinnen und Schüler zu unterstützen.

Es wurde bereits darauf hingewiesen, dass Lernen als ein individueller, aktiver Prozess verstanden wird. Bei der Gestaltung von Lernumgebungen, die sich an den genannten Forderungen orientiert, müssen in hohem Maße das Vorwissen und die Vorerfahrungen der Lernenden durch entsprechende instruktionale Hilfestellungen berücksichtigt werden (Prechtel & Nerdel 2002). Eine Konfrontation mit komplexen Lernsituationen kann je nach Vorwissenstand zu einer Verunsicherung oder sogar Überforderung der Lernenden führen, was dem eigentlichen Ansinnen, der Initiierung eines selbständigen Lernprozesses, entgegenwirkt bzw. ihn verhindert.

Auf Probleme, die authentische und komplexe Lernsituationen mit sich bringen, weisen u. a. Hartinger, Fölling-Albers, Lankes, Marenbach & Molfenter (2001), Klauer (1999) und Stark, Gruber & Mandl (1998) hin. Harp & Mayer (1998) machen darauf aufmerksam, dass bei komplex gestalteten Lernumgebungen die Gefahr besteht, dass sich die Lernenden in so genannte „seductive details“ verlieren und ihnen dabei der Blick für das Wesentliche abhanden kommt (Harp et al. 1998). Bei der unterrichtlichen Umsetzung von Kontexteinheiten ist es daher u. a. zentral, dass vor der eigentlichen Bearbeitung einzelner Aspekte, der Kontext gegliedert und in kleinere Themeneinheiten zerlegt wird.

Situierte und konstruktivistische Lernumgebungen bieten dem Lernenden zahlreiche Möglichkeiten des interessierten, motivierten und selbstwirksamen Lernens, „um dieses Angebot sinnvoll nutzen zu können, bedarf es jedoch bestimmter Kompetenzen, die nicht bei allen Lernenden vorausgesetzt werden können“ (Stark et al. 1998, S. 213). Diese Kompetenzen haben zum einen fachlichen Ursprung (zur Domänenspezifität von Kompetenzen siehe Abschnitt 2.1.3.4 der vorliegenden Arbeit), zum anderen aber auch übergeordneten Charakter, wie eine Darstellung der Grundlagen des selbstgesteuerten Lernens im Folgenden deutlich macht.

2.1.2 Selbstgesteuertes Lernen

Ein übergeordnetes Ziel von *Chemie im Kontext* auf methodischer Ebene ist der Aufbau und die Förderung der Fähigkeit zum selbstgesteuerten Lernen¹ der Schülerinnen und Schüler. Es kann nicht mehr davon ausgegangen werden, dass während der Schulzeit der notwendige Bildungsstand für das Leben als Erwachsener in ausreichendem Umfang erworben wird. Lernende müssen sich vielmehr darauf einstellen, dass ihre Bildungsbestände einem ständigen Ergänzungs- und Erneuerungsprozess unterliegen (Friedrich 1997, Lehrke 1998).

Neben dem Aufbau eines Basiswissens sollte eine zentrale Zielsetzung von Schule und Unterricht sein, dass Schülerinnen und Schüler lernen, wie man lernt. Es wurde bereits deutlich, dass die Lehrkraft in einem Unterricht nach *Chemie im Kontext* nicht als „Überträger“ oder „Vermittler“ von Wissen auf die Schülerinnen und Schüler fungiert, sondern verstärkt als Moderator von Lernprozessen in Erscheinung tritt. Den Lernenden hingegen wird eine aktivere und damit zugleich auch eine verantwortungsvollere Rolle im Lernprozess zugeteilt. Das wird einerseits durch eine methodische Öffnung des Chemieunterrichts z. B. durch verstärkten Einsatz von Gruppen- und Stationsarbeiten, aber auch durch das Einüben der Präsentation und Diskussion von Ergebnissen forciert. Andererseits ist es dafür natürlich notwendig, den Schülerinnen und Schülern „Handwerkszeug“ zur Verfügung zu stellen, welches sie für die selbständige Auseinandersetzung mit verschiedenen Themenbereichen aus naturwissenschaftlicher Sicht benötigen. Für den Aufbau und die Förderung des selbstgesteuerten Lernens ist der Erwerb bestimmter Strategien erforderlich, wie sie z. B. von Simons (1992) als die fünf wichtigen Fähigkeiten für einen selbständigen Lerner formuliert werden.

Der Lernende muss die Fähigkeit besitzen,

- sein Lernen vorzubereiten,
- die Lernhandlung auszuführen,
- sie zu regulieren,
- (seine) Leistung zu bewerten und
- seine Motivation und Konzentration zu erhalten.

¹ Die Begriffe „selbständig“, „eigenständig“, „selbstreguliert“ werden im Zusammenhang mit Lernen synonym verwendet.

Nach Simons (1992) entsprechen diesen fünf Lernfähigkeiten auf Seiten der Schülerinnen und Schüler fünf Lehrfunktionen auf Seiten der Lehrkraft, sie muss ihrerseits

- das Lernen vorbereiten,
- die Lernhandlungen ausführen

und ihr obliegt

- die Handlungsregulation,
- die Leistungsbewertung,
- die Aufrechterhaltung der Motivation und Konzentration.

Friedrich (1997) beschreibt vier Facetten, die für selbstgesteuertes Lernen insbesondere in der Schule wesentlich sind: die erste Facette der *Informationsverarbeitung* beinhaltet zum einen kognitive Lernstrategien, die sich auf die geistigen Aktivitäten beziehen, die von Schülerinnen und Schülern eingesetzt werden, um mit Informationen umzugehen, und zum anderen Strategien, die dieses Umgehen mit Informationen thematisieren, die so genannten „Kontrollstrategien“ (Strategien zum Konfigurieren, Überwachen und Regulieren). Selbständiges Lernen wird zudem durch *emotional-motivationale Strategien* unterstützt, die dafür sorgen, dass u. a. die Lernbemühungen über einen gewissen Zeitraum aufrecht erhalten werden. Für das Lernen werden verschiedene Ressourcen genutzt, wie z. B. Zeit und Medien, und der effiziente *Umgang mit diesen Ressourcen* stellt somit eine weitere Facette des selbständigen Lernens dar. Als letzter, insbesondere für das Lernen in der Schule zentraler Aspekt wird die *soziale Interaktion* genannt, denn für das Lernen in einer Gruppe müssen geeignete Strategien vorhanden sein bzw. erworben werden. Darüber hinaus bedarf es der Vermittlung von fachspezifischen Kompetenzen, damit Schülerinnen und Schüler die Selbständigkeit im Lernprozess in einem naturwissenschaftlichen Fach wie z. B. Chemie überhaupt übernehmen können. Sie müssen z. B. in der Lage sein, Fragestellungen zu erkennen, die mit Hilfe der Chemie beantwortet werden können, und Ergebnisse adressatengerecht, situations- und fachangemessen zu kommunizieren.

Diese verstärkte Aktivität der Schülerinnen und Schüler kommt u. a. im Einsatz von so genannten „offenen“ Methoden zum Ausdruck. Denn nur wenn im Unterricht selbständiges Lernen geübt und ermöglicht wird und sich der Lehrende immer stärker aus dem Lernprozess zurücknimmt, werden die Lernenden schrittweise die Verantwortung für das eigene Lernen übernehmen (Lehrke 1998, Klieme, Artelt & Stanat 2001).

Es gilt hierbei zu beachten, dass keineswegs eine einzige Methode zur Verwirklichung der Förderung des eigenständigen Lernens geeignet ist, sondern vielmehr die Variation und Vielfalt von großer Bedeutung sind, in der auch „traditionellere“ Methoden wie z. B. der Lehrervortrag oder fragend-entwickelnde Unterrichtssequenzen ihre Berechtigung haben. Eine Methode muss ziel-

orientiert in Hinblick auf den Unterrichtsgegenstand, auf die Lerngruppe und auf das Unterrichtsziel ausgewählt werden (Weinert 2001, Bruder 2003).

Bei der Erarbeitung komplexer, überwiegend neuer Themen kann sich eine gut vorstrukturierte Gruppenarbeit anbieten, während sich zur Wiederholung bereits erarbeiteter oder einfacher Themen auch eine Methode wie z. B. ein Gruppenpuzzle eignet, die sehr viel Verantwortung in die Hände der Schülerinnen und Schüler und zudem den Schwerpunkt auf den methodischen Aspekt legt. Mit den beschriebenen Veränderungen in der Schwerpunktsetzung bei der Unterrichtsgestaltung geht eine Veränderung der Rollen in Hinblick auf das Maß der Aktivität von Lernenden und Lehrenden einher. Während der Lehrkraft eine stärker moderierende Funktion zukommt, müssen die Lernenden ihre rezeptive Haltung aufgeben und eine aktivere und verstärkt gestalterische Rolle im Unterrichtsgeschehen einnehmen (Gerstenmaier et al. 1995).

2.1.3. Naturwissenschaftliche Grundbildung

Im folgenden Abschnitt werden die Ziele, die mit der Konzeption *Chemie im Kontext* auf der inhaltlichen Ebene des Lernens der Schülerinnen und Schüler verfolgt werden, aufgezeigt. Die übergeordnete Zielstellung auf der inhaltlichen Ebene orientiert sich an dem Aufbau einer naturwissenschaftlichen Grundbildung (Scientific Literacy). Aus der zu Grunde gelegten Definition einer Scientific Literacy ergeben sich zunächst allgemeine Anforderungen und Ziele für den naturwissenschaftlichen Unterricht. Daraus lässt sich ein Bearbeitungs- und Erkenntnisgewinnungsprozess konkretisieren, der in einem Unterricht nach *Chemie im Kontext* bei der Bearbeitung eines Kontextthemas idealtypisch durchlaufen wird.

2.1.3.1 Kompetenzen und Kompetenzerfassung

Für die Auseinandersetzung mit den Ausprägungen einer naturwissenschaftlichen Grundbildung ist es einleitend notwendig, den zu Grunde gelegten Kompetenzbegriff vorzustellen und abzugrenzen. Bei dem in den folgenden Darstellungen und Überlegungen verwendeten Kompetenzbegriff wird von der Definition der Kompetenz nach Weinert (2001) ausgegangen. Kompetenzen sind „die bei Individuen verfügbaren oder durch sie erlernten kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösung in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert 2001, S. 27f.). Es wird deutlich, dass der Kompetenzbegriff nach diesem Verständnis über inhaltliches Wissen und kognitive Fähigkeiten hinausgeht.

Zudem wird davon ausgegangen, dass Kompetenzen durch Domänenspezifität gekennzeichnet sind und ebenso wie Wissen situiert erworben werden (vgl. Kap. 2.1.1 dieser Arbeit), d. h. Kompetenzen können „als Leistungsdispositionen in bestimmten Fächern oder ‚Domänen‘“ verstanden werden (BMBF 2007, S. 22). Damit grenzt sich der Kompetenzbegriff von übergeordneten, fächerübergreifenden „Schlüsselqualifikationen“ (im Englischen Cross-Curricular Competencies) ab, wie sie z. B. als allgemeine Forderungen in Lehrplänen als Methoden-, Selbst-, Sozial- und Sachkompetenz zu finden sind (vgl. z. B. MfBWFuK des Landes S.-H. Lehrplan Chemie S I, 1997).

Nach heutigem Kenntnisstand muss davon ausgegangen werden, dass sich fächerübergreifende Kompetenzen nicht allgemein ohne konkreten Domänenbezug fördern lassen (Klieme et al. 2001). Zudem wird nach Weinert (2001) die Ausprägung einer Kompetenz eines Individuums durch verschiedenste Facetten, wie z. B. Wissen, Verstehen, Können, Erfahrung und Motivation bestimmt.

Aus diesen Überlegungen zum Kompetenzbegriff ergeben sich verschiedene Aspekte für die Konstruktion von Instrumenten zur Erfassung von Kompetenzen (BMBF 2007). Zum einen spielen fachbezogenes Wissen und fachbezogene Fähigkeiten eine große Rolle, so dass das Vorhandensein einer bestimmten Kompetenz nicht isoliert überprüft werden kann, sondern vielmehr eine Kompetenz leistungsbezogen erfasst und gemessen wird. Zum anderen muss sich die Operationalisierung einer Kompetenz auf eine konkrete Anforderungssituation beziehen, da sie eine Verbindung zwischen Wissen und Können darstellt und als die Befähigung zur Bewältigung von Situationen bzw. Aufgaben verstanden wird. Kompetenzen beschreiben nach dieser Auffassung Fähigkeiten, die nicht naturgegeben sind, sondern von den Lernenden beim Bearbeiten konkreter Fragestellungen erworben werden (BMBF 2007, Hammann 2004). Mit diesen Fähigkeiten können Teilgebiete der gesellschaftlichen Wirklichkeit gestaltet werden. Sie besitzen keinen fertigen Charakter, sondern bleiben für Ausdifferenzierung und Verfeinerungen offen und ermöglichen auf diese Weise einen Prozess des selbständigen Lernens (BMBF 2007).

2.1.3.2 Naturwissenschaftliche Grundbildung nach PISA

Das übergeordnete Ziel, das der Konzeption *Chemie im Kontext* inhaltlich zu Grunde liegt, ist der Aufbau einer naturwissenschaftlichen Grundbildung bei Schülerinnen und Schülern als einem wichtigen Aspekt der Allgemeinbildung, die eine auf diesem Grundverständnis basierende Beschäftigung mit naturwissenschaftlichen Fragestellungen und Problemen im Erwachsenenalter ermöglichen kann (Duit, Häußler & Prenzel 2001, Bündler et al. 2003). Das inhaltliche Verständnis von naturwissenschaftlicher Grundbildung orientiert sich an der Definition, auf der auch das

internationale und nationale Rahmenkonzept für die Erfassung von naturwissenschaftlicher Kompetenz in PISA 2003 basiert. Danach wird naturwissenschaftliche Grundbildung als „die Fähigkeit, naturwissenschaftliches Wissen anzuwenden, naturwissenschaftliche Fragen zu erkennen und aus Belegen Schlussfolgerungen zu ziehen, um Entscheidungen zu verstehen und zu treffen, die die natürliche Welt und durch menschlichen Handeln an ihr vorgenommenen Veränderungen betreffen“ (Deutsches PISA-Konsortium 2002) verstanden.

Im Rahmen dieser Definition wird von drei Aspekten der naturwissenschaftlichen Grundbildung ausgegangen: den naturwissenschaftlichen Prozessen, naturwissenschaftlichen Konzepten und Situationen (Deutsches PISA-Konsortium 2002). An dieser Stelle wird nur auf die naturwissenschaftlichen Prozesse, mit denen sich naturwissenschaftliche Grundbildung beschreiben lässt, genauer eingegangen, für die beiden weiteren Aspekte wird auf entsprechende Veröffentlichungen verwiesen (z. B. Prenzel, Häußler, Rost & Senkbeil 2002, Deutsches PISA-Konsortium 2002, Hammann 2004).

Im Rahmen der internationalen PISA-Rahmenkonzeption 2003 wird von den folgenden fünf Prozessen ausgegangen (Deutsches PISA-Konsortium 2002):

1. naturwissenschaftliche Fragestellungen erkennen
2. naturwissenschaftliche Nachweise identifizieren
3. Schlussfolgerungen ziehen und bewerten
4. gültige Schlussfolgerungen kommunizieren
5. Verständnis naturwissenschaftlicher Konzepte zeigen

Während die ersten vier Prozesse in erster Linie als mentale Prozesse verstanden werden müssen, steht beim fünften Prozess das naturwissenschaftliche Wissen im Zentrum. Da die mentalen Prozesse aber nur auf einen bestimmten Inhalt angewendet und somit gezeigt werden können, wird ein Verständnis naturwissenschaftlicher Konzepte folglich in den ersten vier Prozessen ebenfalls benötigt. Die beschriebene naturwissenschaftliche Grundbildung kann in Kompetenzniveaus näher charakterisiert werden. Als Orientierungsrahmen dienen hierfür die Kompetenzstufen, die Bybee (2002) als Stufen einer Scientific Literacy formuliert hat (zu einer ausführlicheren Beschreibung der Stufen von Scientific Literacy Bybee 2002, siehe auch Duit et al. 2001, am Beispiel des Experimentierens konkretisiert bei Hammann 2004).

Im Zuge der Auswertung der PISA-Erhebung aus dem Jahr 2000 wird von der Existenz einer Dimension ausgegangen, die insgesamt die Entwicklung von naturwissenschaftlicher Kompetenz abbildet. Diese Entwicklung reicht von einem relativ einfachen, naturwissenschaftlichen Verständnis und von der Fähigkeit, eine „relativ leicht begreifbare naturwissenschaftliche Grundbil-

dung zu benutzen“, bis hin zu der Fähigkeit „kognitiv anspruchsvollere Konzepte anzuwenden“ (Deutsches PISA-Konsortium 2002).

Im Rahmen der nationalen Ergänzung zur internationalen PISA-Erhebung 2003 in Deutschland ist die Definition der naturwissenschaftlichen Grundbildung weiter differenziert worden. Insbesondere fand eine Untergliederung in die drei Schulfächer Biologie, Chemie und Physik statt, für die so genannte „Big ideas“ (für die Chemie z. B. „Verbrennung“ und „chemische Verbindungen und Aggregatzustände“) formuliert wurden. Diese „Basic concepts“ oder „Big ideas“ sind nicht mit den Basiskonzepten von *Chemie im Kontext* gleichzusetzen, da die „Big ideas“ nicht darauf angelegt sind, eine gesamte Wissenschaftsdisziplin abzubilden. Zudem stellen sie nicht in allen naturwissenschaftlichen Fächern allgemein anwendbare Konzepte dar, sondern beinhalten z. T. auch speziellere Konzepte, die sich nicht auf andere Fächer übertragen lassen (Deutsches PISA-Konsortium 2002). Während das Konzept der chemischen Bindung auch in anderen naturwissenschaftlichen Fächern zur Anwendung kommen kann, stellt z. B. das Verbrennungskonzept kein übergeordnetes und damit übertragbares Basiskonzept dar.

Darüber hinaus wurden für die nationale PISA-Zusatzerhebung 2003 sieben kognitive Aufgabenanforderungen herausgearbeitet, die für naturwissenschaftliches Denken und Arbeiten charakteristisch sind. Es handelt sich bei diesen kognitiven Anforderungen u. a. um das divergente und konvergente Denken, den Umgang mit (mental) Modellen und das Verbalisieren von Sachverhalten (dazu genauer Kap. 4.1.2 der vorliegenden Arbeit).

2.1.3.3. Nature of Science

Ein weiterer Aspekt der angestrebten naturwissenschaftlichen Grundbildung ist das Bewusstmachen der Besonderheiten, Möglichkeiten und Grenzen der Methoden und Verfahren zur naturwissenschaftlichen Erkenntnisgewinnung („Nature of Science“). Inhaltlich wird der Begriff einer „Nature of Science“ z. B. im Rahmen des „Project 2061 – Science for All Americans“ gefüllt, das sich zum Ziel gesetzt hat, Schülerinnen und Schüler auf die Zukunft vorzubereiten, sie für die technologischen, wissenschaftlichen und gesellschaftlichen Veränderungen und Neuerungen zu öffnen und sie „science literate“ zu machen (American Association for the Advancement of Science (AAAS) 1989). Es werden Inhalte und Fähigkeiten festgelegt, die die nachfolgende Generation wissen und können muss, um auf die Zukunft vorbereitet zu sein und sich in der sich verändernden Welt zurechtfinden zu können.

Im Rahmen dieses Projektes werden verschiedene Facetten einer Science Literacy postuliert, von denen als ein Aspekt die „Nature of Science“ beschrieben wird, die sich aus dem spezifischen Blick der Naturwissenschaften auf die Welt, der charakteristischen Erkenntnisgewinnung der

Naturwissenschaften und der Besonderheiten von naturwissenschaftlichen Untersuchungen und Methoden zusammensetzt (AAAS 2001).

Für jeden Bereich liegen so genannte „Benchmarks“ (AAAS 1993) vor, anhand derer sich der jeweilige Lernstand, z. B. in Bezug auf die drei Dimensionen einer „Nature of Science“, diagnostizieren lässt. In knapper Form zeigt Tabelle 1 beispielhafte Ausprägungen der oben genannten drei Bereiche der „Nature of Science“ (AAAS 1993), wobei die Zusammenstellung Ausprägungen aus verschiedenen Klassenstufen (Levels) enthält:

Dimensionen einer „Nature of Science“	Benchmarks
Scientific World View (naturwissenschaftliche Sicht der Dinge)	z. B. Kenntnis... ...über die potentielle Modifizierbarkeit von Theorien (während einige sehr alte Kenntnisse und Theorien immer noch Gültigkeit haben, sind andere Theorien widerlegt oder überarbeitet worden); ...darüber, dass die Naturwissenschaften alleine keine moralischen Entscheidung treffen können; ...über die Reproduzierbarkeit von Ergebnissen bei konstanten Bedingungen; ...darüber, dass Erkenntnisse und Theorien einem ständigen Prozess des Überprüfens, Überarbeitens und Verwerfens mit dem Ziel des besseren Verstehens unterliegen.
Scientific Inquiry (naturwissenschaftliche Erkenntnisgewinnung, Verfahren und Methoden):	z. B. Kenntnis... ...über die zentrale Rolle der Kommunikation (bei der Ausarbeitung und Diskussion des Vorgehens im Team, bei der Wiedergabe von Beobachtungen, bei der Vorstellung und Diskussion von Ergebnissen usw.); ...über Instrumente und Hilfsmittel, die weitere Informationen liefern können; ...über verschiedene Aspekte einer Untersuchung (Beweissammlung, logische Schlussfolgerung, aber auch Kreativität und Einfallsreichtum); ...darüber, dass Untersuchungsergebnisse zumeist neue Fragen generieren.
Scientific Enterprises	z. B. Kenntnis... ...darüber, dass Arbeit in Gruppen hilfreich ist; ...über die unterschiedlichen Wege und Instrumente der Informationsbeschaffung; ...über die historischen und kulturellen Wurzeln naturwissenschaftlicher Theorien und Konzepte; ...über die Nutzung von Computern; ...über die Rolle der Naturwissenschaften bei gesellschaftlichen Fragen; ...über die Berücksichtigung ethischer und moralischer Aspekte; ...über die zentrale Rolle der Kommunikation (z. B. bei der Entscheidungsfindung in Team usw.); ...darüber, dass jeder naturwissenschaftlich arbeiten kann.

Tab. 1: Zusammenstellung der Dimensionen einer „Nature of science“ mit exemplarischer Nennung von zugeordneten Benchmarks

Dieses Projekt zeichnet sich u. a. insbesondere dadurch aus, dass die Exklusivität von Naturwissenschaften explizit zurückgestellt wird. Vielmehr ist jeder Einzelne dazu in der Lage, naturwissenschaftliches Denken und Handeln nachzuvollziehen und auszuführen, um sich mit gesellschaftlichen Fragestellungen aus naturwissenschaftlicher Sicht auseinander zu setzen.

2.1.3.4 Bildungsstandards im Fach Chemie

Durch die Diskussion um die Einführung verbindlicher nationaler Bildungsstandards für die Fächer an allgemeinbildenden Schulen in Deutschland erfahren Kompetenzen neben dem Wissen als explizites Ziel von Unterricht zusätzliche Aktualität. Während sich die Ausführungen in den beiden ersten Abschnitten zur naturwissenschaftlichen Grundbildung (PISA, „Nature of Science“) auf die Naturwissenschaften insgesamt beziehen und die Chemie insbesondere bei PISA nur eine eher untergeordnete Rolle spielt, werden in den Standards die Anforderungen speziell für die Chemie konkretisiert, daher werden sie im Rahmen dieser Arbeit genauer betrachtet.

Für das Fach Chemie sind Regelstandards für den Mittleren Schulabschluss formuliert, die sich in der Ausprägung der Anforderungen ebenfalls an dem Kompetenzbegriff nach Weinert (2001) orientieren (Kultusministerkonferenz (KMK) 2004a). Für das Fach Chemie sind in Anlehnung an die Einheitlichen Prüfungsanforderungen im Abitur (EPA, KMK 2004b) vier Kompetenzbereiche formuliert worden, wobei diese Bereiche entsprechend den Vorschlägen der EPA jeweils in drei Anforderungsniveaus unterschieden werden (KMK 2004b).

Die vier Kompetenzbereiche, durch die sich nach Auffassung der KMK naturwissenschaftliche Grundbildung beschreiben lässt, sind Fachwissen, Erkenntnisgewinnung, Kommunikation und Bewertung. Ein naturwissenschaftlich gebildeter Lernender kennt und erkennt demnach chemische Phänomene, Gesetzmäßigkeiten und Begriffe und ist in der Lage, sie chemischen Konzepten zuzuweisen; zur Erkenntnisgewinnung setzt er geeignete Methoden ein und versteht es, Modelle zu nutzen; er besitzt die Fähigkeit, sich relevante, fachliche Informationen zu verschaffen, Inhalte zu erarbeiten und Informationen, Inhalte und Ergebnisse zu kommunizieren.

Abschließend kann er die chemischen Facetten eines Themas erkennen und bewerten. Insgesamt müssen die Lernenden mit einem Mittleren Schulabschluss im Sinne der Regelstandards Kompetenzen zeigen, die einerseits auf einer fachinhaltlichen, andererseits auf einer handlungsbasierten Dimension beruhen, denn nur durch beide Dimensionen wird chemisches Arbeiten umfassend beschrieben (KMK 2004a).

Zur näheren Erläuterung zeigt Tabelle 2 beispielhaft Ausprägungen der vier Kompetenzbereiche, die im Fach Chemie erwartet werden (KMK 2004a, S. 10ff.), dabei werden in dieser Darstellung keine Niveauzuordnungen vorgenommen.

Kompetenzbereich	Beispiele für Ausprägungen
Fachwissen	<ul style="list-style-type: none"> • Schülerinnen und Schüler sind in der Lage, den submikroskopischen Bau einzelner Stoffe zu beschreiben. (Aspekte des Stoff-Teilchen-Konzepts) • Schülerinnen und Schüler können aus den Eigenschaften der Stoffe auf ihre Verwendungsmöglichkeiten und auf damit verbundene Vor- und Nachteile schließen. (Aspekte des Struktur-Eigenschafts-Konzepts) • Schülerinnen und Schüler besitzen die Fähigkeit, chemische Reaktionen in Hinblick auf die Stoff- und Energieumwandlung zu beschreiben. (Aspekte des Konzepts zur chemischen Reaktion) • Schülerinnen und Schüler können energetische Erscheinungen bei chemischen Reaktionen auf die Umwandlung eines Teils der in den Stoffen gespeicherten Energie in andere Energieformen zurückführen. (Aspekte des Energie-Konzepts)
Erkenntnisgewinn	<ul style="list-style-type: none"> • Schülerinnen und Schüler sind in der Lage, bei Untersuchungen (v. a. in chemischen Experimenten) relevante Daten zu erheben oder sie zu recherchieren. • Sie besitzen die Fähigkeit, geeignete Modelle zu nutzen, um chemischen Fragestellungen zu bearbeiten.
Kommunikation	<ul style="list-style-type: none"> • Schülerinnen und Schüler hinterfragen die fachliche Richtigkeit von Darstellungen in den Medien. • Sie können chemische Sachverhalte unter der Verwendung von Fachsprache und/oder unter Zuhilfenahme von Modellen und Darstellungen beschreiben, veranschaulichen oder erklären.
Bewertung	<ul style="list-style-type: none"> • Schülerinnen und Schüler besitzen die Fähigkeit, aktuelle, lebensweltbezogene Fragestellungen zu entwickeln, die durch Nutzung fachwissenschaftlicher Erkenntnisse der Chemie einseitig werden und beantwortet werden können. • Sie können chemische Sachverhalte in übergeordnete Problemzusammenhänge einbinden, Lösungsstrategien entwickeln und anwenden.

Tab. 2: Kompetenzbereiche und exemplarische Darstellung möglicher Ausprägungen.

2.1.4 Zusammenfassung

Sowohl die inhaltlichen Auffassungen von naturwissenschaftlicher Grundbildung nach PISA und der vier Kompetenzbereiche der Standards im Fach Chemie als auch das grundlegende Verständnis einer „Nature of Science“, bilden den Ausgangspunkt für die in der Konzeption *Chemie im Kontext* formulierten Arbeitsschritte im Unterricht und die damit angestrebte Kompetenzentwicklung bei der Be- und Erarbeitung eines Kontextes.

Alle drei vorgestellten Ansätze haben zum Ziel, die Naturwissenschaften als zentralen Teil der Bildung eines Menschen aufzuzeigen, indem sie die Besonderheiten der Erkenntnisgewinnung, Methoden und Verfahren in den Naturwissenschaften betonen. Es geht in allen Ansätzen darum, die Inhalte und Fähigkeiten festzulegen, nach denen ein Mensch als naturwissenschaftlich gebildet angesehen werden kann.

Zusammengefasst lässt sich festhalten, dass ein nach diesem Verständnis naturwissenschaftlich gebildeter Mensch in der Lage sein muss, naturwissenschaftliche Fragestellungen als solche zu erkennen. Er muss um die Besonderheiten und Grenzen von Verfahren und Untersuchungsmethoden wissen, mit denen naturwissenschaftlichen Fragestellungen nachgegangen werden kann. Er benötigt ein gewisses Grundrepertoire an fachlichem Wissen und ein Grundverständnis von naturwissenschaftlichen Konzepten, das es ihm ermöglicht, Entscheidungen nachzuvollziehen, zu verstehen und zu bewerten. Nicht zuletzt muss der Mensch die Fähigkeit besitzen, seine

Kenntnisse, Gedanken, Meinungen, Entscheidungen und Schlussfolgerungen adressatengerecht, fach- und situationsangemessen zu kommunizieren.

Die beschriebenen Ansätze zur Beschreibung von naturwissenschaftlicher Bildung besitzen somit gemeinsame Komponenten wie z. B. einen ausgeprägten kommunikativen Aspekt. Zudem spielt in allen Ansätzen der Aspekt des Bewertens eine große Rolle. Darüber hinaus stellen sie alle das Besondere des naturwissenschaftlichen Arbeitens und der Erkenntnisgewinnung heraus, und nicht zuletzt ist ein fundiertes Verständnis zentraler Konzepte unabdingbar. Auf diese Anforderungen muss der naturwissenschaftliche Unterricht die Schülerinnen und Schüler vorbereiten, um sie in der beschriebenen Weise naturwissenschaftlich gebildet aus der Schule zu entlassen. Dem Erreichen dieser Zielstellung hat sich auch ein Unterricht verschrieben, der sich an *Chemie im Kontext* orientiert.

2.2 Der Prozess der Erkenntnisgewinnung in einem Unterricht nach *Chemie im Kontext*

Im Folgenden wird der bereits erwähnte Erarbeitungs- und Erkenntnisprozess vorgestellt, der in einem Unterricht nach *Chemie im Kontext* idealtypisch bei der Bearbeitung eines Kontextes durchlaufen wird. Auf diese Weise lassen sich die genannten Kompetenzbereiche weiter konkretisieren, indem aufgezeigt wird, was der Lernende bei jedem einzelnen Schritt in diesem Prozess leisten muss, d. h. welche kognitiven Fähigkeiten an welchen Stellen im Prozess besonders gefordert sind. Dieser Kreislauf stellt eine Möglichkeit der Beschreibung eines naturwissenschaftlichen Erkenntnisgewinns dar und zeigt auf, welche Kompetenzbereiche und Kompetenzen von zentraler Bedeutung sind. In seiner strukturellen Grundanlage zeigt diese Beschreibung vergleichbare Schritte wie das forschend-entwickelnde Unterrichtsverfahren (Schmidkunz & Lindemann 1976), das sich u. a. durch eine gute Strukturierung des naturwissenschaftlichen Unterrichts auszeichnet (Pfeifer, Lutz & Bader 2002) und Chemielehrkräften wohlbekannt ist.

Im ersten Schritt bei der Auseinandersetzung mit einem neuen Thema ist es nötig, den Themenbereich (Kontext) zu erschließen, indem dazu z. B. Schlagwörter gesammelt und in einer Mindmap strukturiert oder Fragen formuliert werden, die von Interesse sind. Auf diese Weise werden unterschiedliche Bereiche und Aspekte eines komplexen Themas erfasst. In einem nächsten Schritt ist es dann notwendig, aus diesen formulierten Fragen und Themenaspekten die Bereiche zu identifizieren, zu denen die Chemie bzw. der Chemieunterricht bei der Erarbeitung einen Beitrag leisten kann („naturwissenschaftliche Fragen erkennen“). Die Bearbeitung anderer Fragen,

die von der Chemie wegführen, erhalten keine so starke Gewichtung, sie können aber Anknüpfungspunkte für einen fächerübergreifenden Unterricht liefern. Auf die große Bedeutung des Fragerenkennens und des Fragenformulierens als zentralen Aspekt des naturwissenschaftlichen Arbeitens weist z. B. Stäudel (2003a) hin.

Nachdem naturwissenschaftliche Fragen identifiziert worden sind, muss das weitere Vorgehen geplant werden, d. h. es wird festgelegt, was untersucht werden soll, wie die untergeordneten Fragestellungen und ggf. die Hypothesen lauten und welche Methoden zum Einsatz kommen (z. B. Recherche in Büchern, im Internet, Durchführung von Experimenten). Daraufhin folgt die Durchführung der geplanten Untersuchungen, wobei die Ergebnisse dieser Untersuchungen auf geeignete Weise zusammengetragen, dargestellt und kommuniziert werden müssen.

Im Zusammenhang mit naturwissenschaftlichen Denk- und Arbeitsweisen stellt das Experiment (natürlich) den Kern einer naturwissenschaftlichen Untersuchung dar (z. B. Stäudel 2003a, Pfeifer et al. 2002, Hammann 2004). Im Rahmen der vorliegenden Arbeit bleibt diese besondere Stellung des Experimentes bei der naturwissenschaftlichen Erkenntnisgewinnung jedoch unberücksichtigt, vielmehr wird auf spezielle Arbeiten zur Einbindung des Experimentes im Chemieunterricht verwiesen (Pfeifer 2003, Pfeifer et al. 2002).

Die gewonnenen Ergebnisse müssen in einem nächsten Schritt in zweierlei Hinsicht eingeordnet werden: zum einen müssen die Ergebnisse vor dem Hintergrund der Ausgangsfragestellung bewertet (Einordnung in den Kontext), zum anderen müssen sie in einen fachchemischen Zusammenhang gebracht werden (vgl. dazu auch Stäudel 2003a). Dieser zuletzt genannte Schritt steht in einem engen Zusammenhang mit dem abschließenden, übergeordneten Schritt der Abstraktion der Ergebnisse zu übergeordneten Prinzipien. Die gewonnenen Erkenntnisse werden den Basis-konzepten der Chemie zugeordnet, um auf diese Weise ein strukturiertes Wissensfundament zu entwickeln, das durch ständiges Anwenden und eine fortlaufende Erweiterung bereits enthaltener und durch Hinzufügen neuer Inhalte die Grundlage für ein anwendbares Wissen schafft.

Die folgende Abbildung 2 zeigt den idealtypischen Verlauf dieser Erkenntnisgewinnung.

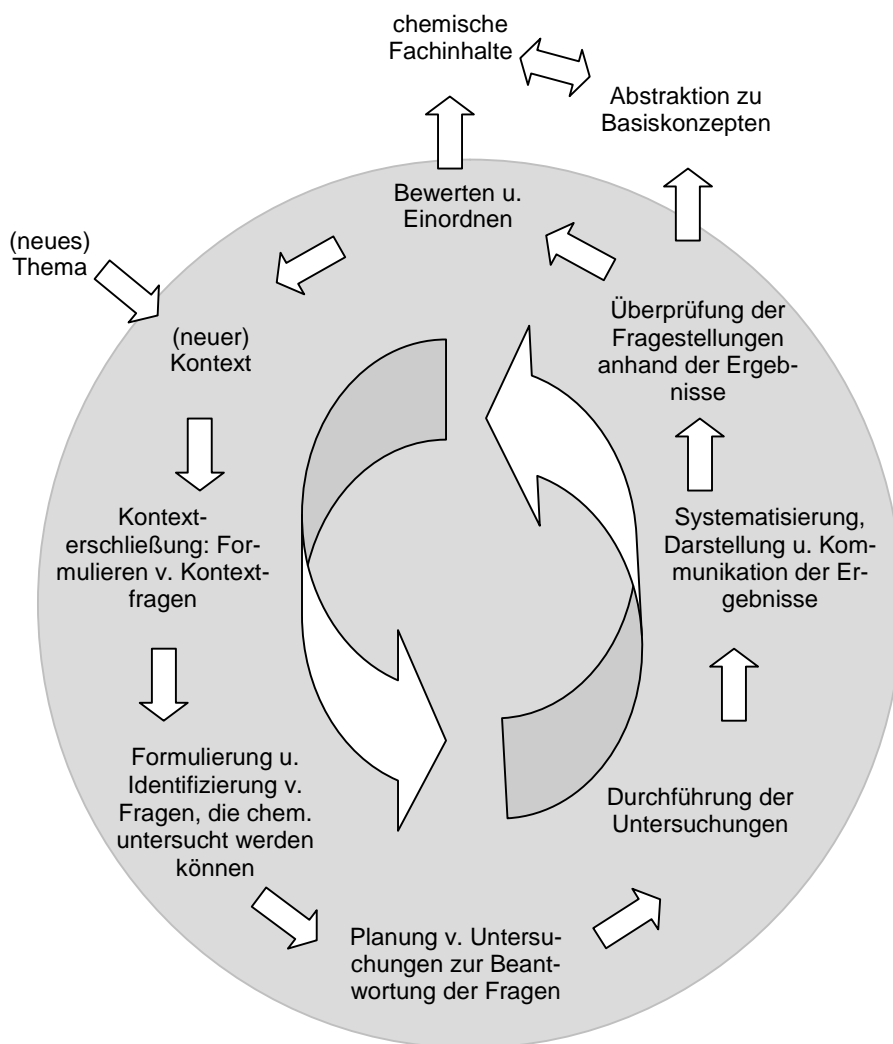


Abb. 2: Idealtypischer Verlauf der Erkenntnisgewinnung: Die graue Hinterlegung beschreibt den Kontext, der für die gesamte Einheit das inhaltliche Gerüst liefert; dieser Kontext wird nur verlassen, um die erarbeiteten Fachinhalte einerseits in einen fachchemischen Zusammenhang einzuordnen und andererseits zu Basiskonzepten zu abstrahieren.

Im Kapitel 4.1.1 der vorliegenden Arbeit wird dieser Prozess erneut aufgegriffen und anhand eines Beispiels aus dem Chemieunterricht veranschaulicht (Tab. 10, S. 57).

Die im Kapitel 2.1.3 aufgezeigten Anforderungen, die an einen naturwissenschaftlich gebildeten Menschen gestellt werden, finden sich in diesem Prozess wieder. Von den drei im Rahmen der internationalen Definition von naturwissenschaftlicher Bildung formulierten Aspekte Prozesse, Konzepte und Situationen lässt sich ein Bogen zur kontextbasierten Unterrichtskonzeption *Chemie im Kontext* schlagen: den Aspekt der Situationen spiegeln die tragenden Kontexte der Einheiten wieder, die Basiskonzepte als strukturierendes Gerüst für die Systematisierung des Wissens werden im Aspekt der Konzepte aufgegriffen, nur die naturwissenschaftlichen Prozesse haben in der Umsetzung der Konzeption bisher zu wenig Beachtung erfahren.

Im Folgenden wird verdeutlicht, wie die beschriebenen Prozesse in einem Unterricht nach *Chemie im Kontext* differenziert berücksichtigt werden können. Bei der Ausdifferenzierung der Definition von naturwissenschaftlicher Grundbildung in der Festlegung des internationalen und nationalen Konzepts der Naturwissenschaften bei PISA lassen sich Zuordnungen zu den formulierten Arbeitsschritten im Erkenntnisprozess vornehmen (Tab. 3), wobei berücksichtigt werden muss, dass die Zuordnungen Gewichtungen und Schwerpunktsetzungen darstellen und die Grenzen dieser Zuordnungen daher fließend sind.

Arbeitsschritte in einem Unterricht nach <i>Chemie im Kontext</i>	PISA – Rahmenkonzept Naturwissenschaften
Kontext erschließen, Fragen zum Kontext formulieren	
Fragen identifizieren, die chemisch untersucht werden können	naturwissenschaftliche Fragestellungen erkennen
Vorgehensweise überlegen und ggf. Experimente oder andere Untersuchungen planen	naturwissenschaftliche Nachweise identifizieren
Experimente oder andere Untersuchungen durchführen	
Ergebnisse darstellen und kommunizieren, ggf. Hypothesen überprüfen	Schlussfolgerungen kommunizieren
Ergebnisse bewerten und in den Kontext einordnen	aus Belegen Schlussfolgerungen ziehen und bewerten,
	Wissen anwenden
Ergebnisse bewerten und in fachchemischen Zusammenhang einordnen	
Einzelergebnisse zu übergeordneten Prinzipien abstrahieren	naturwissenschaftliche Konzepte verstehen

Tab. 3: Gegenüberstellung der Arbeitsschritte im Unterrichtsgang nach *Chemie im Kontext* und der im PISA-Rahmenkonzept formulierten naturwissenschaftlichen Prozesse

Diesen einzelnen Schritten lassen sich wiederum die im Rahmen der nationalen PISA-Zusatzerhebung 2003 formulierten kognitiven Aufgabenanforderungen zuordnen, z. B. stellt die Fähigkeit des divergenten Denkens einen Schwerpunkt bei der Erschließung eines Themenbereichs und bei der zunächst ungerichteten Formulierung von Fragen zum Kontext dar, während z. B. die Fähigkeit des konvergenten Denkens bei der Identifizierung von Fragen, die sich mit Hilfe der Chemie bearbeiten lassen, eine bedeutende Rolle spielt (dazu genauer Kap. 4.2).

Die vier Kompetenzbereiche der Standards für den Chemieunterricht Fachwissen, Erkenntnisgewinnung, Kommunikation und Bewertung (KMK 2004a) fließen ebenfalls in den hier vorgestellten Prozess der Erkenntnisgewinnung ein. Die einzelnen Arbeitsschritte beschreiben Fähigkeiten, die ein Lernender besitzt oder erwerben kann, während er sich im Prozess der Erkenntnisgewinnung bewegt.

Da von einer Domänenspezifität von Kompetenzen ausgegangen wird (BMBF 2007), ist in dem gesamten Prozess die Komponente *Fachwissen* immanent. Einerseits ist der Aufbau von Fachwissen natürlich ein Ziel des Erarbeitungsprozesses, andererseits stellt der Kompetenzbereich Fachwissen aber auch eine Grundlage für die Bearbeitung der einzelnen Schritte dar.

Der Bereich *Erkenntnisgewinnung* wird letztlich durch den gesamten Prozess beschrieben und stellt somit in dieser Darstellung eine übergeordnete Dimension dar.

Auch der Kompetenzbereich *Kommunikation* findet sich in allen aufgezeigten Schritten wieder, erhält aber besondere Bedeutung bei der Darstellung und Diskussion der Ergebnisse und bei der Bewertung und Einordnung der erarbeiteten Inhalte in den Kontext.

Der letztgenannte Bereich der *Bewertung* findet sich überwiegend im fortgeschrittenen Bearbeitungsprozess wieder, wenn es darum geht, die erarbeiteten Inhalte in Hinblick auf die ursprüngliche Fragestellung, aber auch auf ihre Bedeutung im gesamten Kontext hin zu bewerten.

Durch eine „Nature of Science“ mit ihren drei ausdifferenzierten Bereichen wird ebenfalls eine übergeordnete Facette des hier dargestellten Bearbeitungsprozesses beschrieben, denn im Rahmen des Prozesses werden insgesamt Möglichkeiten, Grenzen und Methoden der naturwissenschaftlichen Erkenntnisgewinnung erfahren.

Zusammenfassend orientieren sich die Ziele eines Unterrichts nach *Chemie im Kontext* an den Forderungen nach naturwissenschaftlicher Grundbildung, wie sie im Rahmen von PISA und der Formulierung von Standards diskutiert werden. Die einzelnen Aspekte, die diese Grundbildung charakterisieren, sind in den vorangegangenen Abschnitten vorgestellt worden. Insgesamt wurde ein Erarbeitungsgang aufgezeigt, der in einem Unterricht nach *Chemie im Kontext* zur Bearbeitung eines Kontextthemas idealtypisch durchlaufen wird. Diese Schritte leiten sich aus allgemeinen Forderungen des Verständnisses von naturwissenschaftlichen Grundbildung und einer „Nature of Science“ ab und stellen im Rahmen der vorliegenden Arbeit die theoretische Grundlage für die Konstruktion von Testaufgaben für Kompetenzerfassung dar (siehe Kap. 3.5).

2.3 Wege in die Schulpraxis: Implementation der Konzeption in den Chemieunterricht

Im Anschluss an die Vorstellung der Grundlagen der Konzeption *Chemie im Kontext* folgt in diesem Abschnitt eine verkürzte Darstellung der theoretischen Überlegungen und Forderungen dazu, wie neue Konzepte in die Schule gelangen und in der Praxis umgesetzt werden können. Zudem wird beschrieben, wie diese Realisierung von Seiten der fachdidaktischen Forschung wissenschaftlich unterstützt und begleitet werden kann. Das hier vorgestellte, vom Bundesministerium für Bildung und Forschung (BMBF) und den beteiligten Ländern geförderte Implementationsprojekt (Laufzeit Mai 2002 – Juli 2005) wird am Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN) in Kiel in Kooperation mit den Universitäten Dortmund, Oldenburg und Saarbrücken (bis Sept. 2004)/ Wuppertal (seit Okt. 2004) durchgeführt. In der letzten Phase des Projektes nahmen insgesamt etwa 200 Lehrkräfte aus 12 Bundesländern (Baden-Württemberg, Bayern, Berlin, Brandenburg, Hamburg, Hessen, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, Saarland, Sachsen und Schleswig-Holstein) mit ihren Klassen teil.

2.3.1 Symbiotische Implementationsstrategie

In diesem Projekt wird eine Implementationsstrategie am Beispiel der Konzeption *Chemie im Kontext* erprobt, in der ein symbiotischer Ansatz verfolgt wird.

Bei einer symbiotischen Implementationsstrategie werden, anders als bei so genannten „Top-Down“-Strategien, die „Beteiligten“, in diesem Fall die Lehrkräfte, direkt in den Entwicklungs- und Umsetzungsprozess eingebunden. Während bei Top-Down-Strategien Neuerungen „von oben“ vorgegeben werden (z. B. durch die Einführung nationaler Bildungsstandards oder neuer Lehrpläne), kommt den beteiligten Lehrkräften im Projekt *Chemie im Kontext* ein bedeutender Anteil der Entwicklungs- und Erprobungsarbeit zu. Die verfolgte Strategie orientiert sich an dem Ansatz der „learning communities“ (Putnam & Borko 2000, Zech, Gause-Vega, Bray, Secules & Goldman 2000, Gräsel et al. 2004): In so genannten Schulsets (s. u.) werden Arbeitsgruppen aus Lehrkräften, Vertreterinnen und Vertretern von landeseigenen Fortbildungsinstitutionen und der fachdidaktischen Forschung gebildet.

Durch diese Zusammensetzung wird Sorge dafür getragen, dass einerseits die Lehrkräfte mit ihren persönlichen Fragestellungen und Bedürfnissen den Ausgangspunkt der gemeinsamen Arbeit darstellen, und andererseits auf diese Weise landestypische und in geringerem Maße auch schultypische Besonderheiten bereits in die konzeptionelle Arbeit einfließen und bei der Umsetzung der konzeptionellen Grundlagen in der Praxis berücksichtigt werden können (1. Zwischen-

bericht (2003) und Schlussbericht (2005) zum Forschungsprojekt „Optimierung von Implementationsstrategien bei innovativen Unterrichtskonzeptionen am Beispiel von *Chemie im Kontext*“). Durch die Einbindung übergeordneter Landesinstitutionen der Lehrerfortbildung werden bereits in den Anfängen der Arbeit die Strukturen für die weitere Verbreitung der neuen Ideen in den Ländern rechtzeitig gelegt. Darüber hinaus werden die Lehrergruppen in ihrer Arbeit durch Mitglieder der Projektgruppe begleitet, die inhaltliche Anregungen für die Arbeit geben, die Kommunikation innerhalb des Projektes zwischen Schule und Wissenschaft sicherstellen und den Austausch über Ländergrenzen hinweg fördern.

Bei *Chemie im Kontext* handelt es sich nicht um ein abgeschlossenes Curriculum, das in fertiger Form implementiert werden soll, sondern vielmehr müssen entsprechend den eigenen Bedürfnissen und Schwerpunkten vorhandene Unterrichtseinheiten überarbeitet und angepasst oder neue Einheiten entwickelt werden. Die verfolgte Implementationsstrategie fußt damit auf drei zentralen Aspekten: Kooperation, System- und Prozessorientierung (Schlussbericht 2005).

Um den Fortbestand der Konzeption *Chemie im Kontext* in der Schulpraxis über die Laufzeit des Projektes hinaus sicherzustellen, ist die Phase der Dissemination bereits in der Implementationsphase angelegt: nach einem Jahr wurden in fast allen Bundesländern neue Arbeitsgruppen gegründet, die von Seiten der Projektgruppe nicht mehr in der gleichen Intensität begleitet wurden. Diese neuen Gruppen profitieren insofern von den Arbeiten und Erfahrungen der ersten Arbeitsgruppen, als diese neuen Gruppen von einer konzeptionserfahrenen Lehrkraft aus dem ersten Set geleitet werden. Alle Setlehrkräfte sind zudem dazu angehalten, z. B. an ihrer eigenen oder einer benachbarten Schule multiplikatorisch tätig zu werden. Darüber hinaus schließt sich an das laufende Projekt ein Transfer-Projekt *Chemie im Kontext* an, in dem einerseits die Bildung von Strukturen für die weitere Verbreitung und feste Verankerung der Konzeption in z. B. bestehenden Fortbildungsstrukturen der Ländern vorangetrieben und andererseits Schulsets in neuen Ländern gegründet werden.

Übergeordnetes Ziel des gesamten Projektes ist es, Einflussfaktoren auf den Implementationserfolg zu identifizieren, denn es wird davon ausgegangen, dass sich diese Ergebnisse auch in anderen Projekten nutzen lassen, in denen es darum geht, wissenschaftliche Erkenntnisse und Neuerungen in die Schulpraxis einzuführen. Erste Ergebnisse zu dieser übergeordneten Fragestellung werden bei Fey et al. (2004) berichtet.

2.3.2 Schulsets – Zusammensetzung und Aufgaben

Es wurde bereits beschrieben, dass in den teilnehmenden Ländern so genannte Schulsets gebildet werden. In den Arbeitsgruppen dieser Schulsets werden mit unterschiedlichen Schwerpunkten Unterrichtseinheiten, in denen die Konzeption *Chemie im Kontext* realisiert wird, entwickelt, erprobt, die Umsetzung reflektiert und die Einheiten auf der Basis der gewonnenen Erkenntnisse und Erfahrungen überarbeitet. Die an diesen Arbeitsgruppen beteiligten Personen vertreten vielfältige Perspektiven und Einflussfaktoren auf Schule und Bildung, so dass diese von vornherein in der Arbeit berücksichtigt werden können.

Eine Arbeitsgruppe setzt sich aus zehn bis zwölf Chemielehrkräften, von denen nach Möglichkeit immer zwei aus einer Schule stammen, einem Vertreter einer Fortbildungs- oder administrativen Institution des jeweiligen Landes sowie einem Mitglied der Projektgruppe, folglich aus dem Bereich der fachdidaktischen Forschung, zusammen.

Abbildung 3 stellt dar, wie sich die Arbeit eines Schulsets gestaltet: Neben der Ausarbeitung und Erprobung von Unterrichtsbeispielen stellt die Reflexion des eigenen Unterrichts einen bedeutenden Schwerpunkt der Arbeit dar.

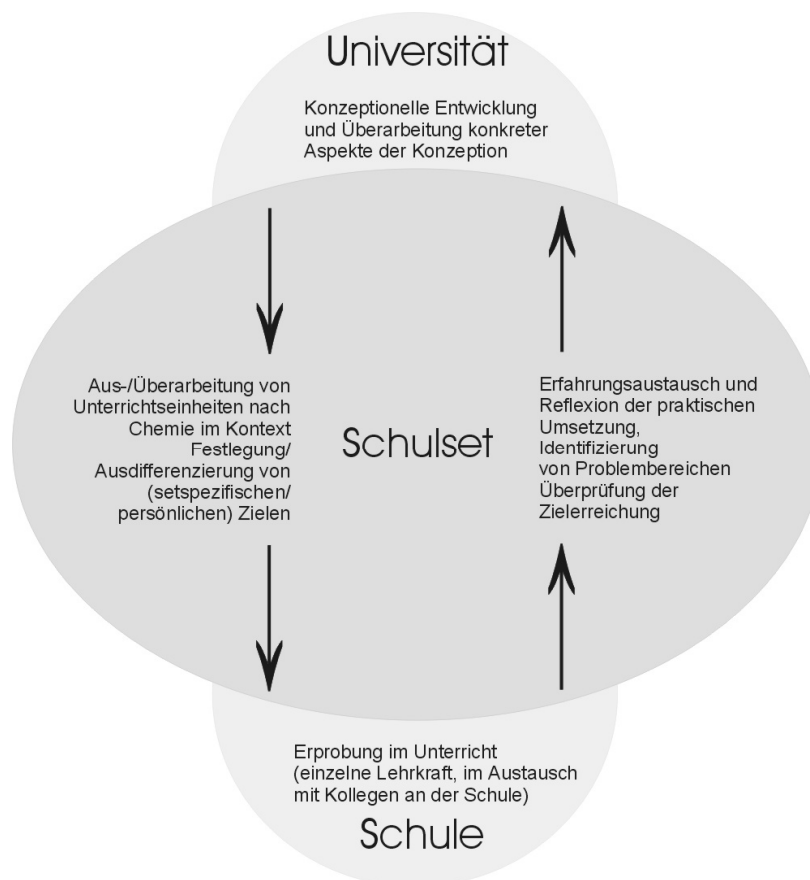


Abb. 3: Aufgaben der Mitglieder eines Schulsets.

Im Rahmen der vorliegenden Arbeit wurde in Zusammenarbeit mit den Lehrkräften der Schulsets in Rheinland-Pfalz und Hamburg Unterrichtseinheiten entwickelt, die mit den Lehrkräften diskutiert und anschließend in der Schulpraxis erprobt wurden. Aus diesen Vorarbeiten ergaben sich erste Erfahrungen bei der Formulierung von Aufgaben und Arbeitsaufträgen in Orientierung an den neuen Schwerpunkten und Zielen des Chemieunterrichts. Der Arbeit sind ausgearbeitete Unterrichtseinheiten als CD (Anhang K) beigelegt, alle aufgeführten Einheiten sind unter Mitwirkung der Autorin entstanden (überwiegend Eigenentwicklung: „Treibstoffe“, Klasse 11; „Chemischer Obstgarten“, Klasse 11/12 (Rheinland-Pfalz); Mitarbeit: „Sauer macht sauber“, Klasse 10 (Hamburg); „Alkohol“, Klasse 11 (Rheinland-Pfalz), Treibstoffe, Klasse 11 (Hamburg)).

2.4 Erfahrungen bei der Umsetzung der Konzeption in der Praxis

Insgesamt lässt sich hinsichtlich der praktischen Umsetzung festhalten, dass sich die eingesetzte Implementationsstrategie im Rahmen des Projektes bewährt hat: Der Unterricht hat sich sowohl in der Wahrnehmung der teilnehmenden Lehrkräfte als auch der Schülerinnen und Schüler verändert. Insbesondere lässt sich eine größere Vielfalt der eingesetzten Methoden und eine stärkere Orientierung an Kontexten im Unterricht feststellen, als durchgehend sehr positiv und produktiv wird dabei die intensive Zusammenarbeit in den Schulsets bewertet (zu Ergebnissen siehe Schlussbericht des Projektes 2005). In den drei ersten Projektjahren haben sich in den arbeitenden Lehrergruppen aber auch verschiedenste Probleme und Fragen bei der praktischen Umsetzung der Konzeption gezeigt, die sich zum Teil auf landestypische oder gruppenspezifische Strukturen zurückführen lassen (Ausstattung der Schulen, Entlastung der teilnehmenden Lehrkräfte, gruppenspezifische Prozesse) (Lang & Bündler 2004, Jäger 2004). Es lassen sich darüber hinaus aber auch Problembereiche ausmachen, die in nahezu allen Schulsets aufgetreten sind und die in der Konzeption *Chemie im Kontext* selbst begründet liegen. Diese Problembereiche liefern insbesondere Ansatzpunkte für die Schwerpunktsetzungen in einzelnen Arbeitsgruppen, aber auch für die konzeptionelle Arbeit in der Projektgruppe, so erhielt z. B. im Rahmen der Setarbeit die Frage nach der Einführung und Umsetzung der Basiskonzepte im Unterricht neue Aktualität (Bündler et al. 2003). Im Folgenden werden zwei übergeordnete Problemstellungen aufgezeigt, die auch für die vorliegende Arbeit Anknüpfungspunkte darstellen.

Problembereich 1: „Unübersichtlichkeit“

Ein Unterricht nach *Chemie im Kontext* zeichnet sich u. a. dadurch aus, dass vielfältigste Gestaltungsmittel und Methoden zum Einsatz kommen, die in erster Linie die Selbständigkeit der Ler-

nenden im Unterrichts- und Lernprozess fördern sollen. Diese Forderung geht v. a. mit einer (vermehrten) Umsetzung von „offenen Methoden“ im Chemieunterricht einher (dazu auch Kap. 2.1 der vorliegenden Arbeit). Darüber hinaus werden die Schülerinnen und Schüler im Planungsprozess für den Unterricht (in erhöhtem Maße) eingebunden, und sie nehmen damit Einfluss darauf, welche inhaltlichen Schwerpunkte im Unterricht gesetzt werden. Die Öffnung des Unterrichts einerseits und die Aufgabe des Planungsmonopols der Lehrkraft andererseits führen für die Lehrenden, aber auch für die Lernenden zunächst zu einer gewissen „Unübersichtlichkeit“ des Unterrichtsverlaufs in Hinblick auf Lernwege und Lehr-Lern-Ziele des Unterrichts. Abbildung 4 verdeutlicht die von vielen Lehrkräften und auch zahlreichen Schülerinnen und Schülern anfänglich empfundene „Unübersichtlichkeit“ eines Unterrichts, der sich an den Grundlagen der Konzeption *Chemie im Kontext* orientiert.

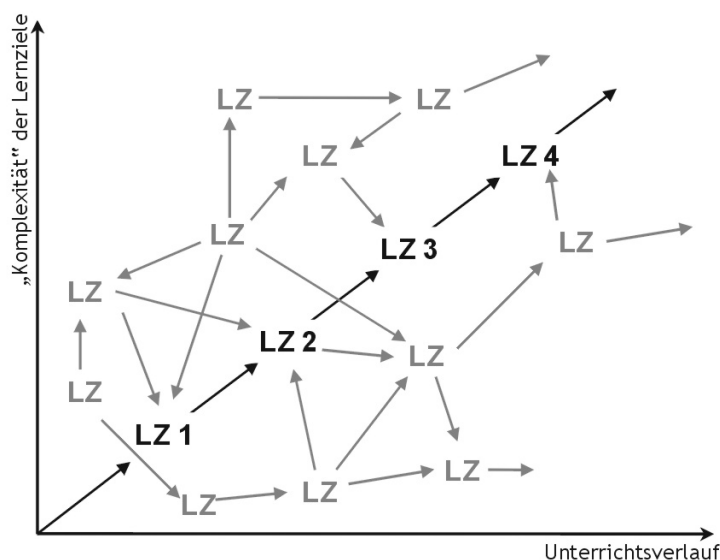


Abb. 4: Gegenüberstellung eines fragend-entwickelnden Unterrichts und einem Unterricht nach *Chemie im Kontext* hinsichtlich der Lernziele; Lernziele des fragend-entwickelnder Unterricht = LZ, sich aus den Kontexten ergebende Lernziele → LZ; (verändert nach Ellenberger 1993, nach Suwelack im Rahmen eines Sattreffens Rheinland-Pfalz).

In einem offenen, stärker durch die Schülerinnen und Schüler strukturierten Unterricht ist es mindestens ebenso wichtig wie in einem fragend-entwickelnden Unterricht, dass auch die Schülerinnen und Schüler die Ziele des Unterrichts klar vor Augen haben und ihnen strukturierende Hilfsmittel angeboten werden, die den „roten Faden“ des Unterrichts für die Lernenden deutlich machen (z. B. eine Mindmap, in der die Fragestellungen festgehalten sind, oder das Sicherstellen der Präsenz der Kontextfrage, an deren Beantwortung gerade gearbeitet wird). Ausgehend von den Theorien des situierten Lernens erwerben Schülerinnen und Schüler ihr Wissen mehr oder minder „zufällig“ (Bünder et al. 2003) und dieses zufällige Wissen hat zunächst keinen systematischen Bezug zu bereits vorhandenen oder noch zu erwerbenden Wissensbeständen. Als Hilfs-

mittel dienen auf der Ebene einer Einheit eine Kontextleitfrage und, ebenfalls in den Einheiten, aber auch vor allem übergeordnet, die Basiskonzepte. Im Verlauf der Bearbeitung einer Themenstellung, z. B. im Rahmen einer Stationsarbeit, muss den Lernenden die thematische Anbindung an den Kontext ständig präsent und das Ziel ihrer Arbeit zu jedem Zeitpunkt bewusst sein. Die Lehrkräfte stehen durch die Vielfalt der Lernwege und Lernziele vor dem Problem, einschätzen zu müssen, wo ihre Schülerinnen und Schüler zu einem gegebenen Zeitpunkt stehen. Es bedarf daher Instrumente, die in der Lage sind, der Lehrkraft relativ einfach und schnell einen Überblick zu verschaffen, wo Schülerinnen und Schüler oder Lerngruppen im Bearbeitungsprozess stehen. Vielversprechende Hilfsmittel stellen prozessbegleitende Instrumente wie z. B. Concept Maps (Stracke 2004), Mind Maps oder Lernbegleitbögen (Parchmann et al. 2003, Schmidt et al. 2003) dar, deren Funktionalität und Tragfähigkeit aber z. T. noch überprüft werden muss. Problembereich 1 kann somit zusammenfassend als „Unübersichtlichkeit der Lernwege und Lernziele“ beschrieben werden.

Problembereich 2: „Leistungskontrolle“

Die Tatsache, dass nicht mehr von einer streng aufeinander aufbauenden Abfolge von durch die Lehrkraft festgelegten Lernzielen ausgegangen werden kann, sondern eben eine Fülle von z. T. von den Lernenden selbst festgelegten Lernzielen und -wegen an diese Stelle tritt, muss natürlich bei der Lernzielkontrolle und Leistungsbeurteilung berücksichtigt werden. Das Überprüfen erreichter Lernziele gestaltet sich dann als schwierig, wenn der Fülle und den unterschiedlichen Ausrichtungen möglicher Lernziele Rechnung getragen werden soll. In einem Unterricht nach *Chemie im Kontext* erhalten zusätzlich zu den Lernzielen die Lernwege eine größere Bedeutung, d. h. es bedarf zudem Leistungskontrollen, die in der Lage sind, etwas über Lernwege, folglich über das, was die Schülerinnen und Schüler „neben der Chemie“ lernen, auszusagen.

Erworbene Kompetenzen erhalten neben dem Erwerb von fachlichem Wissen große Bedeutung, denn die Schülerinnen und Schüler sollen in die Lage versetzt werden, sich bestimmte Themengebiete aus naturwissenschaftlicher Sicht selbständig zugänglich zu machen und zu bearbeiten (vgl. Kap. 2.2). In diesem Zusammenhang stellt sich die Frage, was sich überhaupt im Rahmen von Leistungskontrollen abprüfen lässt und welche Instrumente dafür geeignet erscheinen. Insgesamt rücken in einem Unterricht nach *Chemie im Kontext* neben den Lernzielen die Lernwege und neben dem erworbenen Wissen die erworbenen Kompetenzen in den Vordergrund. Problembereich 2 umreißt somit die Frage nach geeigneten Formen der Lernziel- und Leistungsüberprüfungen.

2.5 Neuer Unterricht – Alte Bewertungsmaßstäbe?!

Aus den beiden Problembereichen lassen sich verschiedene Aspekte ableiten, die in der vorliegenden Arbeit genauer ausgeführt werden. So sollten z. B. Aufgaben einer schriftlichen Leistungskontrolle in einem Unterricht nach *Chemie im Kontext* idealerweise alle Facetten des vorangegangenen Unterrichts in angemessenem Maße aufgreifen. Lernende ermessen u. a. anhand der Aufgaben aus Leistungsüberprüfungen, „worauf es im Unterricht ankommt“ (Duit et al. 2001, Häußler & Lind 1998, von Davier & Hansen 1998): Werden die Leistungen von Schülerinnen und Schülern in einem Unterricht nach *Chemie im Kontext* mit „herkömmlichen“ Testaufgaben mit einem deutlich ausgeprägten fachchemischen Schwerpunkt gemessen, wird den Lernenden auf diese Weise vermittelt, dass eben doch „die Chemie“ das eigentlich Wichtige und das „Drumherum“ für die Note wenig von Bedeutung ist.

Zusammenfassend lässt sich festgehalten, dass es für eine veränderte Form des Chemieunterrichts neuer und anderer Instrumente bedarf, die in der Lage sind, Lehrenden wie Lernenden aufzuzeigen, wo sie im Lernprozess stehen. Es sind Instrumente und Verfahren nötig, die Leistungen so messen, dass dem vorangegangenen Unterricht in möglichst zahlreichen Aspekten Rechnung getragen wird. Das bedeutet u. a., dass an die Seite des gemessenen Fachwissens als weitere Komponente das Erfassen der Kompetenzen tritt, die im naturwissenschaftlichen Unterricht von zentraler Bedeutung sind. Zu diesem Zweck muss nach einer detaillierten Zielklärung für den Chemieunterricht eine bewusste Auseinandersetzung mit Aufgaben dahingehend erfolgen, ob diese Aufgaben in der Lage sind, den Kompetenzerwerb einerseits zu fördern und zu begleiten, andererseits aber auch zu kontrollieren und zu messen.

3 Aufgaben – Gestaltung und Einsatz im naturwissenschaftlichen Unterricht

Aufgaben nehmen derzeit in der aktuellen fachdidaktischen, aber auch in der bildungspolitischen Diskussion einen bedeutenden Stand ein, nicht zuletzt angeregt durch die Leistungsvergleiche im Rahmen von PISA, in denen die wenigen Aufgaben, die der Öffentlichkeit zugänglich gemacht sind, verdeutlichen, mit welchen Anforderungen die Lernenden in diesen Leistungsmessungen konfrontiert werden (Stäudel 2004). Die Ergebnisse aus flächendeckenden Leistungsstudien wie PISA, aber auch z. B. die Einführung von Vergleichsarbeiten, nationalen Bildungsstandards und Zentralabitur machen eine Auseinandersetzung mit den Themengebieten „Aufgaben“ und „Aufgabekultur“ in den Fachdidaktiken, aber auch und vor allem in der Schule notwendig.

Die Expertengruppe des Gutachtens zur Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts (BLK 1997) schreibt Aufgaben eine große Bedeutung dahingehend zu, dass sie dazu dienen können, bei Schülerinnen und Schülern Lernen zu motivieren, ein verständnisvolles Erschließen von Inhalten zu ermöglichen, Inhalte und Fähigkeiten zu üben und Wissen zu konsolidieren. Aus diesem Grund liefert die Beschäftigung mit (Weiter-)Entwicklung von Aufgabenstellungen und -gestaltungen zahlreiche Anknüpfungspunkte für eine Veränderung und Verbesserung des mathematisch-naturwissenschaftlichen Unterrichts (BLK 1997).

Im folgenden Abschnitt wird ein kurzer Abriss der Diskussion zu diesem Thema gegeben. Dabei werden ausgehend von Aufgaben für den naturwissenschaftlichen Unterricht insgesamt fachspezifische Besonderheiten der Aufgabenformulierung für den Chemieunterricht herausgestellt.

Die beiden Perspektiven, aus denen das Themengebiet Aufgaben betrachtet werden kann, sind, wie bereits erwähnt, zum einen die Forschungsperspektive, zum anderen die Perspektive des Unterrichts. Eine Aufgabenbearbeitung durch Lernende gibt Lehrkräften oder Forschenden Aufschluss z. B. über Lernstände oder Leistungen der Schülerinnen und Schüler (Forschungsperspektive mit entsprechender Fragestellung); die Aufgabenformulierung selbst kann aber auch etwas über den tatsächlichen Unterricht aussagen, in dem sie eingesetzt wurde (Perspektive des Unterrichts).

Zunächst wird auf die Bedeutung von Aufgaben im Unterricht eingegangen, denn der Einsatzbereich von Aufgaben ist vielfältig: Sie dienen einerseits als diagnostisches Instrument bei der Lernziel- und Leistungskontrolle bei Schülerinnen und Schülern (Aufgaben als diagnostisches Instrument oder aber als Instrument zur Bewertung der Leistungen der Lernenden), andererseits sind sie strukturierendes Element des Unterrichts. In Abhängigkeit von dem Ziel, welches mit einer Aufgabe verfolgt wird, gibt es Unterschiede in den Aufgaben z. B. in der Art ihrer Gestal-

tung. Eine Aufgabe kann aber auch verschiedene Funktionen erfüllen, die von ihrer jeweiligen Einbindung in den Unterrichtsverlauf abhängig sind.

Die Art der im Unterricht eingesetzten Aufgaben kann Aufschluss über den eigentlichen Unterricht geben, z. B. lassen sich anhand einer Analyse von Aufgaben Einseitigkeiten in den Anforderungen feststellen (Aufgaben als Gestaltungs- und diagnostisches Werkzeug von Unterricht).

Im Anschluss daran werden Aufgaben in Hinblick auf ihre Gestaltung und Klassifizierung eingehender betrachtet. In der Literatur finden sich verschiedene Ansätze zur Systematisierung von Aufgaben, die in diesem Kapitel in knapper Form dargestellt werden. Abschließend werden Hinweise und Erkenntnisse zur konkreten Konstruktion von Testaufgaben zur Kompetenzerfassung aus der Literatur zusammengefasst.

3.1. Aufgaben und Aufgabenkultur – Begriffsfestlegungen

Nach Häußler et al. (1998a) werden Aufgaben als definierte Probleme verstanden, für die es (mindestens) eine Lösung gibt und die sich in verhältnismäßig kurzer Zeit bearbeiten lassen. Nach dieser Definition werden Aufgaben ausgeschlossen, die sich auf das Üben von (physikalischen) Arbeitsweisen (z. B. Schülerexperimente) und nicht in erster Linie auf das Anwenden von und Umgehen mit (physikalischen) Gesetzen und Begriffen beziehen. Ausgeschlossen werden durch diese Definition ebenfalls Aufgaben, die sich durch einen komplexen Inhalt und eine offene Fragestellung auszeichnen (diese Aufgabenform wird als Problem bezeichnet).

Das Verständnis des Aufgabenbegriffs, welches dieser Arbeit zu Grunde gelegt wird, ist weiter gefasst. In einer umfassenderen Definition des Aufgabenbegriffs wird von einer Aufgabe als „Aufforderung zum Lernhandeln“ ausgegangen (Bruder 2003). Eine Aufgabe als Aufforderung zum Lernhandeln präsentiert sich dem Bearbeitenden mit drei Komponenten, d. h. eine Aufgabe besitzt eine Ausgangs- und Endsituation und Lösungswege („Transformationen“), die die Verbindung zwischen Ausgangs- und Endsituation herstellen (Bruder 2003). Diese Definition des Aufgabenbegriffs lehnt sich damit eng an die Problem-Definition aus der Denk- und Problemlösepsychologie an (Dörner 1975), nach der ein Problem durch einen unerwünschten Ausgangs-, einen erwünschten Endzustand und eine Barriere, die das unmittelbare Umwandeln vom ersten in den zweiten Zustand verhindert, gekennzeichnet ist.

Durch Auslassungen und Besetzungen der drei von Bruder (2003) beschriebenen Positionen lassen sich verschiedene Aufgabentypen konstruieren und charakterisieren. Für den mathematisch-naturwissenschaftlichen Bereich sind z. B. Begründungs- und Beweisaufgaben typisch: Ausgangs- und Endsituation sind inhaltlich besetzt und die Lernenden sind aufgefordert, den Lösungsweg zu formulieren, der die Transformation von der einen Situation zur anderen erklärt.

Durch das Besetzen bzw. Freilassen der drei genannten Komponenten einer Aufgabe in verschiedenen Kombinationen entwirft Bruder (2003) ein Zielstrukturmuster, nach dem sich Aufgaben analysieren lassen (Aufgabentypen mit Beispielen aus der Mathematik Bruder 2003, weitere Beispiele aus der Mathematik Büchter & Leuders 2005).

Die der vorliegenden Arbeit zu Grunde gelegten Definition des Aufgabenbegriffs beschränkt sich nicht auf Aufgaben im engeren Sinne, sondern umfasst auch den Bereich der Probleme. Eine Aufgabe wird im Folgenden allgemein als Aufforderung zum Lernhandeln verstanden, und umfasst damit alle schriftlichen und mündlichen Arbeitsaufträge und Aufgaben, die z. B. im Unterricht eingesetzt werden.

Der Begriff der Aufgabekultur umfasst die Gesamtheit der im Unterricht realisierten Aufgaben und Arbeitsaufträge und beschreibt zudem die „Orchestrierung“ (Duit et al. 2002) der Aufgaben, ihren Inhalten und Zielen in Zusammenhang mit ihrer Einbindung in das tatsächliche Unterrichtsgeschehen und Anpassung an die Unterrichtsmethode.

3.2 Einsatzgebiete von Aufgaben im Unterricht

Aufgaben können aus zwei Perspektiven betrachtet werden: zum einen aus dem Blickwinkel der Unterrichtsgestaltung und -ziele, zum anderen kann sich der Blick auf die Adressaten der Aufgaben richten, auf die Lernenden. In beiden Fällen ist es erforderlich, sich der Maßstäbe, mit denen Aufgaben betrachtet und untersucht werden und die die Aufgaben ihrerseits messen und untersuchen, bewusst zu sein. Um sich sinnvoll mit dem Themengebiet Aufgaben und Aufgabekultur auseinandersetzen zu können, ist es in einem ersten Schritt notwendig, die genannten Maßstäbe zu formulieren und diese für alle Beteiligten transparent zu machen.

Das Formulieren dieser Maßstäbe für Aufgaben schließt sich insgesamt an die Zielformulierung für den Unterricht an.

3.2.1 Gestaltungsmittel und Diagnosewerkzeug im Unterricht

Aufgaben erhalten entsprechend ihrer Einbindung in das tatsächliche Unterrichtsgeschehen ihre jeweilige Funktion (Häußler et al. 1998a). Das bedeutet, eine Aufgabe kann letztlich erst mit Blick auf den konkreten Unterricht hinsichtlich ihres Zwecks kategorisiert werden. Zudem können Aufgaben prinzipiell für einzelne Unterrichtsphasen konstitutiv sein (Häußler et al. 1998a), z. B. wird einer Erarbeitungsphase eine komplexe Aufgabe vorangestellt oder der Einstieg in ein neues Unterrichtsthema durch eine interessante Frage aus der Erfahrungswelt der Lernenden eingeleitet.

Da Aufgaben ein maßgebliches Gestaltungselement von Unterricht sind, können sie Informationen zu dem konkreten Unterricht liefern und nehmen demnach eine diagnostische Funktion auf Seiten des Lehrenden ein. Im Unterricht eingesetzte Aufgaben können als Reflexionsanlass genutzt werden, um sich mit dem eigenen Unterricht auseinanderzusetzen. Eine Analyse der im Unterricht eingesetzten Aufgaben kann Hinweise darauf liefern, ob in der Art der Aufgabengestaltung eine Einseitigkeit in den Anforderungen oder in der Zweckgebundenheit von Aufgaben vorherrscht. Der überwiegende Einsatz von Aufgaben, die sich durch eine sehr kleinschrittige Aufgabenformulierung auszeichnen, verweist möglicherweise darauf, dass es der Lehrkraft schwer fällt, die Schülerinnen und Schüler im Unterrichtsprozess alleine und eigenverantwortlich arbeiten zu lassen oder andere und damit auch „falsche“ Lösungswege zu ermöglichen.

Eine Analyse von im Unterricht eingesetzter Aufgaben zeigt auf diese Weise Anknüpfungspunkte auf, an denen eine Veränderung von Unterricht ansetzen kann. Übergeordnete Fragestellungen können sich z. B. damit befassen, wann Aufgaben wie zu welchem Zweck in den Unterricht eingebunden werden, wie die eigene Aufgabekultur aussieht, welche Bewertungsmaßstäbe persönlich für wichtig erachtet und welche „von außen“ vorgegeben werden. Ein Instrument für die Untersuchung der eigenen Aufgabekultur beschreibt die „Spinnennetz“-Methode (Stäudel 2003b): die einzelnen „Fäden“ des Netzes kennzeichnen bestimmte selbst festgelegte oder von außen vorgegebene Kategorien (z. B. zur Offenheit einer Aufgabe, zur Einbettung in einen Kontext oder zum eingeschätzten Schwierigkeitsgrad). Auf diesen Netzfäden werden dann Aufgaben z. B. einer Klausur eingeordnet: Je weiter eine Aufgabe am Rand des Netzes eingeordnet wird, desto größer ist die Offenheit der Aufgabe, desto stärker ist ihre Kontextanbindung und desto höher ist ihr Schwierigkeitsgrad. Werden mehrere Aufgaben auf einem Netz eingeordnet, lassen sich auf einen Blick besondere Häufungen und auffällige Verteilungen feststellen. Der Erstellung eines solchen Netzes muss zwangsläufig eine Auseinandersetzung mit den Zielen und Kategorien, an denen die untersuchten Aufgaben gemessen werden, vorausgehen. Die Lehrkraft ist auf diese Weise gezwungen, die Maßstäbe, die sie an Aufgaben anlegt oder die von außen an Aufgaben gelegt werden, zu formulieren und für sich transparent zu machen.

3.2.2 Bewertungsinstrument für Lernende

Neben der bereits vorgestellten, diagnostischen Funktion in Hinblick auf den Unterricht und die Unterrichtsgestaltung dienen Aufgaben ganz maßgeblich als Diagnoseinstrument bei Lernenden zur Feststellung ihres Leistungsstandes (mit Bewertungsfunktion), des Erreichens von Lernzielen und zur Erhebung des Lernstands (ohne Bewertungsfunktion). Die letztgenannten Aufgaben haben ausschließlich eine diagnostische Funktion und besitzen keine bewertenden Aspekte. Bewer-

tungsaufgaben hingegen werden in Leistungskontrollen eingesetzt und haben daher für die Lernenden einen expliziten Bewertungsaspekt.

Diagnose besitzt generell zwei Grundintentionen: einerseits die Diagnose von Lernerfolg, andererseits die Diagnose von Lernschwierigkeiten, diese beiden Ausrichtungen müssen bei der Erstellung und Auswertung von Aufgaben berücksichtigt werden.

Gerade bei Aufgaben, die zu diagnostischen Zwecken formuliert werden, ist es wichtig, sich der Anforderungen und des Anforderungsniveaus bewusst zu sein. Darüber hinaus müssen Aufgaben, die zu bewertenden Zwecken eingesetzt werden, für die Lernenden als solche gekennzeichnet werden. Auf den hier beschriebenen Diagnose- und Bewertungsaspekt von Aufgaben auf Seiten der Lernenden ist bereits an früherer Stelle eingegangen worden (Kap. 2.5).

3.3 Ansätze zur Klassifizierung von Aufgaben zur Analyse und Konstruktion

In der fachdidaktischen Literatur finden sich verschiedene Vorschläge zur Systematisierung von Aufgaben, einige dieser Ansätze werden im Folgenden vorgestellt. Darüber hinaus ist es wichtig festzuhalten, dass es nicht in erster Linie darum geht, neue Aufgabentypen zu entwickeln, sondern für eine Vielfalt an im Unterricht eingesetzten Aufgaben zu sorgen, in der auch bereits Vorhandenes seine Berechtigung besitzt. Ebenso wie bei dem vielfältigen Methodeneinsatz gilt es auch hier, zielorientiert und zweckgebunden Aufgaben auszuwählen und einzusetzen (vgl. Kap. 2.1 dieser Arbeit, Bruder 2003). Die schönste Aufgabe fruchtet nicht, wenn sie nicht in angemessener und geeigneter Weise in den Unterricht eingebunden ist (Duit et al. 2002, Büchter et al. 2005).

3.3.1 Allgemeine Forderungen – Aufgabeneinsatz, Kontextorientierung und Öffnung von Aufgaben

Allgemeine Forderungen in Hinblick auf neue Aufgaben und eine veränderte Aufgabenkultur lauten z. B. Vielfalt in Einsatz und Gestaltung (Duit et al. 2002), inhaltliche Anbindung der Aufgabenstellungen an die Lebenswelt der Lernenden (BLK 1997) und „Öffnung“ von Aufgaben und damit Zulassen verschiedener Lösungswege (Häußler et al. 1998a).

Auf den ersten Aspekt der Aufgabenvielfalt ist bereits im vorangegangenen Abschnitt kurz eingegangen worden. Einerseits geht es um die Entwicklung und das Kreieren neuer Aufgaben, andererseits geht es aber auch um die vielfältigen Einsatzmöglichkeiten von erprobten und auch neuen Aufgaben im Unterrichtsgeschehen. Eine Forderung der Verfasser der SINUS-Expertise lautet, dass Aufgaben generell eine viel größere Bedeutung im eigentlichen Unterricht erhalten müssten, d. h. dass sie vom Rand in das Zentrum des Unterrichtsgeschehens gerückt werden soll-

ten (BLK 1997). Damit geht wiederum eine Veränderung und Erweiterung der im Unterricht eingesetzten Methoden einher, denn nur in einem in Richtung methodischer Vielfalt veränderten Unterricht ist es möglich, eine andere Aufgabenkultur zu realisieren (Duit et al. 2002).

Die zweite Forderung umschreibt das inhaltliche Einbetten der Aufgabenstellungen in die Lebenswelt der Schülerinnen und Schüler, um auf diese Weise die Beschäftigung mit der Aufgabenstellung zu motivieren und für die Lernenden sinnhafter zu gestalten. Mit einer Kontextorientierung der Aufgabenstellungen geht jedoch eine Zunahme der Komplexität der Aufgaben einher, denn eine gesellschaftlich oder persönlich relevante Fragestellung zeichnet sich in der Regel durch ein Zusammenspiel mehrerer Aspekte und Blickwinkel aus.

Im Zusammenhang damit steht die letztgenannte Forderung nach der Öffnung von Aufgaben: Für eine Aufgabe gibt es nicht nur einen korrekten, sondern mehrere mögliche Lösungswege, die für den Unterricht alle von gleicher Bedeutung sind.

Aber es reicht natürlich nicht aus, Aufgaben zu „öffnen“ und in einen lebensweltlichen Kontext einzubetten. Viel zentraler ist es, die Lernenden zunächst schrittweise an die Bearbeitung der Aufgaben herangeführt werden (z. B. durch Differenzierung einer Aufgabe in verschiedene Schwierigkeitsgrade durch Teilaufgaben oder Bereitstellung von Lernhilfen), um einer Überforderung insbesondere der leistungsschwächeren Schülerinnen und Schüler vorzubeugen und auch ihnen Kompetenzerleben zu ermöglichen.

Ziel ist demnach nicht, die Lernenden mit einer Vielfalt von (neuen) Aufgabentypen zu konfrontieren, sondern sie in die Lage zu versetzen, damit „umzugehen“, indem Herangehensweisen und Bearbeitungswege (und nicht nur die Lösungen) angesprochen werden. Häußler et al. (1998a) fordern beispielsweise für die Physik, Aufgaben nicht bereits so zu formulieren, dass sie eine algebraische Lösung nahelegen, sondern bewusst auch andere Lösungswege zu thematisieren (vgl. Kap. 3.3.3.3; verschiedene Beispiele dazu in Häußler et al. 1998a).

Aufgaben lassen sich generell durch eine Oberflächen- und eine Tiefenstruktur beschreiben. Die Forderung nach der Anbindung an und Einbettung von Aufgaben in lebensweltliche Kontexte bezieht sich auf die Gestaltung der Oberflächenstruktur einer Aufgabe (z. B. Sicherheitsmaßnahmen an einer Tankstelle als Aufgabenthema), während die Tiefenstruktur das Prinzip darstellt, dessen Anwendung den Lösungsweg (z. B. die Betrachtung der Modellebene der Teilchen des Benzins) beschreibt. Die Gestaltung der Oberflächenstruktur mit wechselnden Kontexten kann einerseits dazu dienen, bei den Lernenden eine Motivation für die Aufgabenbearbeitung herzustellen, andererseits wird so aber auch ein Beitrag zur Förderung der flexibleren Wissensanwendung geleistet. Die die Aufgaben bearbeitenden Schülerinnen und Schüler müssen hinter der Oberflächenstruktur das Wesentliche, den Kern der Aufgabe erkennen.

3.3.2 „Gute“ Aufgaben

An eine Beschreibung der Merkmale „guter“ Aufgaben schließt sich im folgenden Kapitel 3.3.3 eine Darstellung verschiedener Systematisierungsansätze an. An dieser Stelle werden zunächst Überlegungen vorgestellt, die sich ganz allgemein mit dem Qualitätsmerkmal „gut“ für Aufgaben auseinandersetzen (Langlet & Freiman 2003, Aufschnaiter & Aufschnaiter 2001, Fischer & Draxler 2001). Trotz aller Problematik dieser Beurteilung von Aufgaben wird auf die Ausführungen deshalb eingegangen, da sie einen ersten Überblick zur Betrachtung von Aufgaben vermittelt. Dabei darf aber nicht vergessen werden, dass die Zuweisung „gut“ bei einer Aufgabe nur in Hinblick auf eine konkrete Unterrichtsgestaltung (Büchter et al. 2005), eine konkrete Klasse oder sogar einen einzelnen Lernenden und einen konkreten Unterrichtsinhalt vorgenommen werden kann, mit anderen Worten: eine „gute“ Aufgabe wird nur durch einen „guten“ Unterricht wirklich gut.

Lehrkräfte beurteilen intuitiv, was eine „gute“ oder eine „schlechte“ Aufgabe ist. Sie entwickeln mit der Zeit ein Gefühl dafür, welche Aufgaben zu ihrem Unterricht „passen“ und welche nicht, d. h. sie haben bestimmte Wertmaßstäbe für Aufgaben entwickelt und verinnerlicht, anhand derer sie Aufgaben beurteilen. Fischer et al. (2001) weisen darauf hin, dass es für die deutschsprachige Literatur charakteristisch ist, dass die Bearbeitung und Lösung einer Aufgabe eher erfahrungsbasiert und intuitiv bestimmten Schwierigkeitsniveaus oder erforderlichen Kompetenzen und Fertigkeiten zugewiesen wird.

Lehrkräfte entwickeln, erproben und optimieren ständig Aufgaben für den Unterrichtseinsatz oder die Evaluation ihres Unterrichts in Klausuren und Lernzielkontrollen. Um einen Austausch über Aufgaben und Aufgabenkultur zu ermöglichen, ist es von zentraler Bedeutung, dass die eigenen Wertmaßstäbe transparent gemacht und kommuniziert werden, um das Formulieren und Optimieren von Aufgaben bewusst und kriteriengeleitet zu gestalten.

Langlet et al. (2003) haben Gütekriterien zusammengestellt, nach denen Aufgaben (für den Biologieunterricht) als „gute“ Aufgaben bewertet werden können. In diesem Zusammenhang stellen auch sie heraus, dass es nicht allein auf eine gelungene Gestaltung und Formulierung einer Aufgabe ankommt, sondern dass die Einbindung und Passung zum Unterricht (ziel- und situationsangemessener Aufgabeneinsatz) ausschlaggebend ist, wobei sie allerdings die Richtung der Einflussnahme von Aufgabentypus und Unterrichtsverfahren offen lassen (Langlet et al. 2003). Die formulierten Aspekte einer guten Aufgaben sind in Tabelle 4 zusammengestellt (Langlet et al. 2003).

Durch eine „gute“ Aufgabe...	<ul style="list-style-type: none"> ... wird die Kenntnis von (biologischen) Fakten gefördert, ... wird der Aufbau eines qualitativen Verständnisses von Begriffen unterstützt, ... kann vorhandenes Wissen konsolidiert und Grundwissen aufgebaut werden, ... wird Wissen aus dem ursprünglichen Kontext herausgelöst und auf diese Weise flexibilisiert, ... wird der Umgang mit verschiedenen Repräsentationsformen von Informationen (Diagramme, symbolhafte Darstellungen usw.) gefördert, ... werden Alltagsvorstellungen der Lernenden aufgegriffen, ... werden Bezüge zum Alltag und der Lebenswelt der Lernenden hergestellt.
„Gute“ Aufgaben...	<ul style="list-style-type: none"> ... enthalten kumulative Elemente, ... ermöglichen mehrere Vorgehensweisen, ... zielen nicht auf schematisierte Lösungswege, ... ermöglichen das Formulieren eigener Fragestellungen und Ziele, ... sind präzise formuliert, ... enthalten verabredete Operatoren als Präzisierungshilfe.

Tab. 4: Merkmale einer „guten“ Aufgabe (Langlet et al. 2003).

Aufschnaiter et al. (2001) leiten aus empirischen Untersuchungen zur Bearbeitung von Physikaufgaben Erkenntnisse zur Gestaltung einer neuen Aufgabenkultur ab, wobei sie jedoch darauf hinweisen, dass die Ergebnisse nicht ohne weitere Überprüfung auf die anderen Naturwissenschaften anwendbar sind. Sie unterscheiden bei ihren Aufgabenstellungen eine inhaltliche, eine Abstraktions- und eine zeitliche Dimension: Eine „gute“ Aufgabe zeichnet sich nach ihren Erkenntnissen dadurch aus, dass die Anzahl der inhaltlichen Aspekte, die verknüpft werden können, an die Erfahrungen, die die Lernenden mit dem Aufgabeninhalt haben, angepasst ist. Zudem muss in Hinblick auf den Kenntnisstand der Schülerinnen und Schüler entschieden werden, ob der Inhalt bereits als etwas Abstraktes erfasst werden kann oder ob er noch konkret beschrieben werden muss. Abschließend muss bei der Formulierung einer Aufgabe die zeitliche Komponente berücksichtigt werden: Nach den Ergebnissen von Aufschnaiter et al. (2001) muss die Aufgabe für die Schülerinnen und Schüler innerhalb von fünf Minuten als Ganzes lösbar sein oder in geeignete Teilaufgaben zerlegt und damit vorstrukturiert werden, weil die Bearbeitung der Aufgabe danach abgebrochen wird (zur Bedeutung der Vorstrukturierung von Aufgaben auch Precht et al. 2002).

Die genannten Aspekte machen deutlich, dass die Passung zwischen Aufgabe und Lernendem von zentraler Bedeutung ist, dass die Aufgabenstellung folglich möglichst nah am tatsächlichen Kompetenzniveau des Schülers oder der Schülerin orientiert sein muss. Darüber hinaus weisen die Autoren darauf hin, dass nach ihren Erkenntnissen Schülerinnen und Schüler eine Arbeit in Gruppen häufig als unproduktiv erleben, so dass auch dieser Aspekt in einer guten Aufgabe berücksichtigt werden muss, indem z. B. vorher deutlich gemacht wird, welches Ziel im Vordergrund steht, das Erarbeiten eines neuen Fachinhaltes oder das Einüben sozialer Kompetenzen (Aufschnaiter et al. 2001).

3.3.3 Systematisierung der verschiedenen Ansätze zur Kategorisierung von Aufgaben

Im Folgenden werden ausgewählte Ansätze zur Auseinandersetzung mit Aufgaben und ihrer Gestaltung genauer beleuchtet. Die Überlegungen, die sich in der Literatur finden lassen, beschäftigen sich auf sehr unterschiedlichen Ebenen mit Aufgaben und Aufgabengestaltung. Die Betrachtung kann von rein formalen Kriterien ausgehen, die bei der Gestaltung von Aufgaben eine Rolle spielen, sie kann auf die Art des erfragten Wissens oder die erforderlichen Fähigkeiten bis hin zu einer Zuordnung zu Kompetenzbereichen abzielen. Zudem kann die konkrete Einbindung in das unterrichtliche Geschehen und damit Zweck und Ziel der Aufgabe thematisiert werden. Darüber hinaus wird in einigen Fällen über die erste Systematisierung eine zweite gelegt, z. B. durch die Festlegung von Schwierigkeitsniveaus bzw. Kompetenzstufen.

Zum einen gibt es die bereits vorgestellten allgemeinen Forderungen, nach denen Aufgaben übergeordnet beurteilt und kategorisiert werden können, wie Grad der Offenheit einer Aufgabe, Ausmaß der Kontextorientierung und Art der Unterrichtseinbindung. Zum anderen gibt es Ansätze, die sich durch einen eher formalen Schwerpunkt bei der Analyse und Beurteilung von Aufgaben auszeichnen. Darüber hinaus sind verschiedenen Typisierungen von Aufgaben formuliert und beschrieben worden, von denen ebenfalls einige exemplarisch dargestellt werden. Ein großer Bereich der Auseinandersetzungen mit Aufgaben beschäftigt sich aber nicht mit eher gestalterischen Merkmalen von Aufgaben, sondern setzt den Schwerpunkt auf die Anforderungen, die die Aufgaben in Hinblick auf die zum Bearbeiten erforderlichen Kompetenzen und Kompetenzbereiche an den Lernenden stellen. Diesem letzten Bereich werden u. a. die konzeptuellen Überlegungen zu den Ausarbeitungen von PISA-Aufgaben und der Formulierung der Bildungsstandards Chemie zugewiesen.

Die folgende Darstellung orientiert sich an Kategorien, die von Fischer & Draxler (2001, 2002, vgl. auch Woest 2004) für die Konstruktion und Bewertung von Physikaufgaben in sehr ausführlicher Form ausgearbeitet worden ist.

Eine Zuordnung der verschiedenen Ansätze erfolgt ggf. nach einer Schwerpunktsetzung in den einzelnen Beiträgen.

Die Klassifizierung einer Aufgabe kann erfolgen:

- nach Inhalt, Lehrplan oder Richtlinien
- zu bestimmten Unterrichtsphasen
- nach Art der möglichen Lösungswege
- nach ihrem Antwortformat
- nach dem Grad ihrer Offenheit

- nach dem erforderlichen Experimentierverhalten
- entsprechend ihren Anforderungsmerkmalen
- zu vorgegebenen Kompetenzstufen oder Kompetenzbereichen

Eine andere Herangehensweise an Aufgaben ist das Arbeiten mit Beispielaufgaben, an die Systematisierung schließt sich eine kurze Darstellung dieser besonderen Aufgabenform an.

3.3.3.1 Zuordnung einer Aufgabe nach Inhalt, Lehrplan oder Richtlinien

Eine übergeordnete Einteilung von Aufgaben erfolgt nach ihrem inhaltlichen Schwerpunkt in der Oberflächen- und Tiefenstruktur. In diesem Zusammenhang wird deutlich, ob eine Aufgabe kontextorientiert formuliert ist, an welchen Aspekten aus der Lebenswelt sie anknüpft und welche fachlichen Themen durch sie aufgegriffen werden. Darüber hinaus kann es hilfreich sein, Aufgaben in Lehrpläne oder Rahmenrichtlinien einzuordnen, um so einen schnellen Überblick zu ermöglichen. Wie auch bei den folgenden Kriterien muss für die Zuordnung einer Aufgabe jeweils ein Schwerpunkt festgelegt werden, nach dem die Aufgabe kategorisiert werden kann, denn die Grenzen zwischen verschiedenen fachlichen oder kontextlichen Themenstellungen oder zwischen den zum Bearbeiten der Aufgabe erforderlichen Fähigkeitsanforderungen sind häufig fließend. In diesen Bereich fallen auch die Klassifizierungen einer Aufgabe in Hinblick auf die angestrebte Klassen- oder Altersstufe der Zielgruppe und auf das (naturwissenschaftliche) Fach. Fischer et al. (2002) fordern zudem eine Einschätzung des Bezugs zur Realität der Aufgabenstellung und der damit zusammenhängenden Interessenslage der Schülerinnen und Schüler hinsichtlich des Aufgabeninhalts.

Diese erste Klassifizierung einer Aufgabe ist hilfreich, da die Lehrkraft sich eingehend damit auseinandersetzen muss, ob und wie die Aufgabe ein Kontextthema aufgreift, ob der Kontext nur ein Aufhänger oder für die gesamte Aufgabe tragendes Prinzip ist. Darüber hinaus wird deutlich, ob in der Aufgabe „alte“ Wissensbestände benötigt werden oder ob sich die Bearbeitung mit aktuell erworbenem Wissen bewerkstelligen lässt.

3.3.3.2 Zuordnung einer Aufgabe zu bestimmten Unterrichtsphasen

Eine zweite Form der Klassifizierung schließt an das an, was bereits im Zusammenhang mit Aufgaben als diagnostisches Instrument des Unterrichts gesagt wurde. Eine Aufgabe kann verschiedenen Phasen des Unterrichts zugeordnet werden, Fischer et al. (2002) unterscheiden in diesem Zusammenhang die drei Phasen Erarbeitung, Übung und Leistungsmessung; Tabelle 5 gibt einen kurzen Überblick zu den von Fischer et al. (2002) formulierten Anforderungen an die Aufgabenstellungen in den Phasen.

Unterrichtsphase	Anforderung an Aufgabenstellung
Erarbeitung	Mit den Aufgaben... ... werden aktuell erlernte Inhalte sinnvoll angewendet, ... werden mögliche Einsatzbereiche für dieses neu erworbene Wissen angedeutet, ... wird den Schülerinnen und Schülern Rückmeldungen zu möglichen Verständnisschwierigkeiten gegeben, ... wird nicht zu viel „älteres“ Wissen aufgegriffen.
Übung	Mit den Aufgaben... ... wird Wissenstransfer gefördert, ... wird Motivation über Kompetenzrückmeldung gefördert, ... werden unterschiedliche Lösungswege erlaubt, ... werden erarbeitete Kenntnisse in für die Schülerinnen und Schülern interessanten und authentischen Kontexten angewendet, ... wird die Anwendung „älterer“ Wissensbestände mit neu erlerntem Wissen verknüpft.
Leistungsmessung	In den Aufgaben... ... werden alle verfolgten Lernziele erfasst, ... werden mehrere, in ihrem Anspruch unterschiedliche Lösungswege zugelassen, ... können „ältere“ Wissensbestände aufgegriffen und mit neuen kombiniert werden, wenn es sich um sehr zentrale, grundlegende Lerninhalte handelt, ... muss deutlich ausgewiesen sein, dass sie zur Leistungsmessung eingesetzt werden.

Tab. 5: Zuordnung einzelner Aufgabenanforderungen zu den drei Phasen des Unterrichts (Fischer et al. 2002).

Häußler et al. (1998a) fügen diesen drei Phasen fünf weitere hinzu: die Phase des Einstiegs, der Wiederholung, Individual- und Gruppenphasen und Hausaufgaben, wobei zu berücksichtigen ist, dass bei dieser Kategorisierung verschiedene Ebenen betrachtet werden. Aufgaben für diese hier genannten Phasen zeichnen sich durch unterschiedliche Anforderungen aus, die durch die Aufgabenstellungen erfüllt werden müssen. Während eine Aufgabe, die in einer Übungsphase eingesetzt wird, das Erlernte festigen und den Wissenstransfer ermöglichen soll (Häußler et al. 1998a), ist es bei einer Aufgabe für die Erarbeitungsphase zentral, dass sie das z. B. Verstehen neuer Konzepte befördert (Fischer et al. 2002).

Durch die Zuordnung einer Aufgabe zu einer bestimmten Unterrichtsphase wird eine Zielaussage getroffen, in der die Zweckgebundenheit einer Aufgabe zum Ausdruck kommt. Ob eine Aufgabe für das Erreichen dieses Ziels wirklich geeignet ist, hängt in hohem Maße von der jeweiligen Lerngruppe ab. Eine genauere Analyse anhand der folgenden Ordnungskriterien kann Aufschluss darüber geben, an welchen Stellen es möglicherweise Schwierigkeiten geben könnte.

3.3.3.3 Einordnung einer Aufgabe nach Art der möglichen Lösungswege

Eine zentrale Forderung in Hinblick auf die Formulierung von Aufgaben ist das Zulassen von mehreren Lösungswegen. Vor allem für Physikaufgaben wird festgestellt, dass sie aufgrund ihrer Formulierung häufig eine einzige, algebraische Lösung forcieren (Häußler et al. 1998a). Häußler et al (1998a) schlagen vor, neben den üblichen algebraischen Lösungen insbesondere qualitative, halb-quantitative, graphische oder pragmatisch-unorthodoxe Lösungen anzuregen, um auf diese

Weise nichtalgebraischen Repräsentationen eines gegebenen Sachverhaltes im Unterricht eine größere Bedeutung zu geben.

Fischer et al. (2002) empfehlen eine etwas andere Kategorisierung, die sich vor allem darin unterscheidet, dass das Experiment als zentrales Instrument der Erkenntnisgewinnung in den Naturwissenschaften explizit Raum erhält. Sie regen an, Lösungen den folgenden vier Kategorien zuzuweisen: Experimentelle, halbquantitative, rechnerische und theoretische Lösungen.

Eine von Woest (2004) vorgeschlagene Kategorisierung lehnt sich eng an die Kategorien von Fischer et al. (2002) an, indem er für die Chemie rechnerische, experimentelle und theoretische Lösungen vorschlägt.

Einen für die Chemie typischen Lösungsweg stellt das Formulieren von Reaktionsgleichungen dar, die in einer Kategorisierung berücksichtigt werden müsste. Darüber hinaus sind für den Chemieunterricht sicherlich Unterscheidungen in stoffliche und modellhafte Lösungswege zu treffen, wobei jedoch zu bedenken ist, dass diese Lösungen nicht als qualitativ gleichwertig betrachtet werden können.

3.3.3.4 Klassifizierung einer Aufgabe nach ihrem Antwortformat

Im Zusammenhang mit der Forderung nach mehreren Lösungswegen und damit nach Öffnung der Aufgaben steht die Klassifizierung des Lösungsweges. Bei Aufgaben, die verschiedene Lösungen ermöglichen, müssen alle Lösungswege kategorisiert werden. Fischer et al. (2002) gehen in Anlehnung an die in der TIMS-Studie eingesetzten Antwortformate von drei möglichen Formaten aus: Multiple-Choice-Aufgaben mit einer oder mehreren richtigen Ankreuzmöglichkeiten, Kurzantworten, die von den Schülerinnen und Schülern in Form eines Stichwortes, eines kurzen Satzes, einer Zahl oder einer kurzen Rechnung selbst formuliert werden, und Aufgaben mit einem erweiterten Antwortformat. Diese Lösungen bestehen u. a. aus ausführlichen Rechnungen, Beschreibungen und ggf. Durchführungen von Experimenten, längeren Textantworten, Skizzen oder Diagramme.

Diese Einteilung ist recht grob, es lässt sich eine differenziertere Unterteilung der Antwortformate vornehmen, in dem die für die zweite und dritte Kategorie genannten Beispiele als einzelne Kategorien gefasst werden. Für die Chemie müsste wiederum das Formulieren von Reaktionsgleichungen als eine weitere (untergeordnete) Kategorie des zweiten oder dritten Antwortformats hinzugefügt werden.

3.3.3.5 Klassifizierung einer Aufgabe nach dem Grad ihrer Offenheit

Über die Öffnung von Aufgaben als zentrale Forderung im Umgang mit und in der Gestaltung von Aufgabenstellungen ist bereits berichtet worden, daher ist es sinnvoll, für einzelne Aufgaben den Grad der Offenheit zu beschreiben. Fischer et al. (2002) unterscheiden dabei in einem ersten Schritt, ob mögliche Lösungswege vorgegeben sind. Im Falle einer Lösungswegvorgabe wird die Intensität dieser Vorgabe in drei Stufen kategorisiert (Fischer et al. 2002):

- *Stufe 1:* Die Aufgabe lässt mehrere Lösungswege zu und liefert keinen Hinweis auf einen bestimmten Lösungsweg.
- *Stufe 2:* Für den Lösungsweg sind wiederum mehrere Möglichkeiten zugelassen, von denen einige Varianten explizit thematisiert werden.
- *Stufe 3:* Der Lösungsweg ist durch die Aufgabenstellung vorgegeben.

Bei den Stufen 2 und 3 wird zudem der Grad der Detailliertheit, mit der die Vorgabe formuliert ist, erfasst:

- Der Lösungsweg ist z. B. durch Handlungsanweisungen allgemein vorgegeben.
- Für den Lösungsweg ist bereits eine Methode vorgeschlagen.
- Der Lösungsweg ist in der Aufgabe genau beschrieben.

Eine weitere Aufgabentypisierung anhand ihres Grades der Offenheit in ihrer Gestaltung nimmt Graf (2001) vor. Er unterscheidet offene, halboffene und geschlossene Aufgaben und ordnet dieser Einteilung verschiedene Aufgabentypen zu. Den Grad der Offenheit legt Graf ebenfalls über die Lösung der Aufgabe fest: Bei einer offenen Aufgabe ist die Lösung weder der Lehrkraft noch den Lernenden bekannt (z. B. Gestaltungsaufgaben, Deutungsaufgaben), bei einer halboffenen Aufgabe kennt zumindest die Lehrkraft die erwarteten Antworten (z. B. Freiantwortaufgaben, Ergänzungsaufgaben, Aufbauaufgaben), während bei einer geschlossenen Aufgabe der Lehrkraft die Lösung bekannt ist und den Schülerinnen und Schülern alle Lösungselemente präsentiert werden (z. B. verschiedene Formen von Auswahlaufgaben, Zuordnungsaufgaben).

3.3.3.6 Zuordnung einer Aufgabe nach dem erforderlichen Experimentierverhalten

Der Kategorisierungsansatz von Fischer et al. (2002) zeichnet sich insbesondere dadurch aus, dass er das Experiment ins Blickfeld bei der Auseinandersetzung mit Aufgaben rückt. Daher wird in diesem Ansatz eine Klassifizierung von Aufgaben entsprechend dem angestrebten Experimentierverhalten der Aufgabenbearbeiter vorgenommen:

Dabei wird unterschieden zwischen einem Verhalten beim Experimentieren, das dadurch gekennzeichnet ist, dass die Lernenden eine Versuchsanleitung „nachkochen“ (*imitatorisch*), ohne Anleitung, nur anhand vorgegebener Gerätschaften selbständig Messungen durchgeführt werden (*organisierend*) oder die relevanten Messgrößen herausarbeiten, Hypothesen formulieren und daraus einen Versuchsaufbau ableiten (*konzeptuell*).

Diese Kategorisierung enthält eine qualitative Stufung des Experimentierverhaltens, wobei festzuhalten ist, dass die erste Stufe im naturwissenschaftlichen Unterricht wohl am häufigsten vertreten ist und die letzte Stufe das eigentliche, aber zugleich sehr anspruchsvolle Ziel des Unterrichts darstellt.

3.3.3.7 Klassifizierung einer Aufgabe entsprechend ihren Anforderungsmerkmalen

Einen großen Raum in der Aufgabenkategorisierung nach Fischer et al. (2002) nimmt die Klassifizierung der Anforderungsmerkmale einer Aufgabe ein. Dazu wird ausgehend von den Merkmalen der TIMS-Aufgaben ein Kriterienkatalog formuliert, der sich aus 16 Anforderungsmerkmalen zusammensetzt, die hier nicht im Einzelnen genannt werden (dazu Fischer et al. 2001, 2002). Für jedes der genannten 16 Merkmale wird bei der Kategorisierung einer Aufgabe entschieden, ob es für die Bearbeitung und Lösung einer Aufgabe keine Relevanz besitzt, von Bedeutung ist oder ob sich die Aufgabe ohne dieses Merkmal gar nicht lösen lässt. Auf diese Weise wird nicht nur ein charakteristisches Anforderungsprofil für eine (Teil-) Aufgabe festgelegt, sondern für eine Aufgabe lässt sich abschließend ein Katalog aufstellen, der besagt, welche der 16 Anforderungsmerkmale für die erfolgreiche Bearbeitung der gesamten Aufgabe bedeutungslos, brauchbar oder unabdingbar sind. Diese Art des Vorgehens zeichnet sich dadurch aus, dass für jedes Anforderungsmerkmal jeweils neu entschieden wird, welche Bedeutung es für die Bearbeitung hat. Bei dieser Art der Kategorisierung muss demnach nicht die Entscheidung getroffen werden, ob es in der Aufgabenanforderung z. B. um ein Verständnis für symbolische Zeichnungen oder um das Überwinden von Fehlvorstellungen geht, sondern es ist beides für die Lösung der Aufgabe zentral. Es findet gerade keine Schwerpunktsetzung statt, wie sie z. B. bei der Formulierung von PISA-Aufgaben vorgenommen werden muss (siehe weiter unten in diesem Abschnitt). Dieses von Fischer et al. (2002) beschriebene Vorgehen kann hilfreich sein, wenn man sich die Komplexität einer Aufgaben bewusst machen möchte, es kann hinderlich sein, wenn man sich einen Überblick über bereits kategorisierte Aufgaben verschaffen möchte, eben weil diese Kategorisierungen komplex und sehr umfassend sind.

Ein weiterer theoretischer Ansatz, der im Zusammenhang mit einer Kategorisierung anhand von Aufgabenanforderungen vorgestellt wird, ist die Itemkonstruktion für die deutsche Zusatzerhebung im Rahmen der internationalen PISA-Erhebung 2003.

Die Ergebnisse der nationalen PISA-Erhebung 2000 (Deutsches PISA-Konsortium 2002) sind ausdifferenziert worden, indem insgesamt sieben für das naturwissenschaftliche Denken und Arbeiten charakteristische kognitive Aufgabenanforderungen formuliert und inhaltlich ausgeschärft werden (Deutsches PISA-Konsortium 2002). Für die Teilaufgabe eines PISA-Items gibt es hinsichtlich der Anforderungsmerkmale nur die Entscheidung, ob sie für die Bearbeitung der Aufgabe zentral ist oder nicht, zudem ist nur eine Anforderungsfestlegung pro Teilaufgabe möglich. Diese Form der Klassifizierung anhand von Anforderungsmerkmalen ist bedeutend problematischer, da die Schwerpunktlegung auch immer Interpretation beinhaltet und die Grenzen zwischen den einzelnen Anforderungen oftmals fließend sind.

Diese sieben kognitiven Aufgabenanforderungen sind bereits im Kapitel 2.1.3.2 dieser Arbeit kurz genannt worden, werden aber an dieser Stelle in der nachfolgenden Tabelle 6 ausführlicher erläutert und anhand von Beispielen veranschaulicht (Deutsches PISA-Konsortium 2002):

	Erläuterung	Spezifizierung (ggf. aus chemischer Sicht)	Beispielaufgabe
Aufgabenanforderung	Divergentes Denken beschreibt die Fähigkeit, alternative Antworten auf eine Frage formulieren zu können, die aber in ihrer Summe ein Grundverständnis des zu Grunde liegenden Sachverhaltes ausdrücken	Diese Anforderung zeichnet sich durch einen Kreativitätsaspekt aus, der sich gerade nicht an Fächergrenzen hält, sondern natürlich in erster Linie fachübergreifend beschrieben werden muss: kreatives Aufspüren von Fragestellungen und Zusammenhängen, Vermutungen äußern, Fragen formulieren, Hypothesen bilden, Brainstorming, Gedankenspiele, Perspektivwechsel, ...	<i>An einer Tankstelle in Barcelona erhält man ein Benzin, das eine andere Zusammensetzung hat als das Benzin, das in Hamburg aus der Zapfanlage kommt. Äußert Vermutungen dazu, warum es verschiedene Benzinsorten gibt!</i>
	Konvergentes Denken beschreibt die Fähigkeit, verschiedene Wissensbestände und Informationen aufeinander beziehen zu können	logisches Schlussfolgern, logisches Verknüpfen, hinter Fragen die Chemie entdecken, Untersuchungen aus Fragen und Hypothesen ableiten, Auswertung von Daten und Ergebnissen, Deutung von Ergebnissen	<i>Benzin als Treibstoff, Feuerzeugbenzin, Reinigungsbenzin – die Bezeichnung „Benzin“ ist allen Produkten gemeinsam. Sammelt die gemeinsamen Eigenschaften, die für Benzine charakteristisch sind!</i>
	Sachverhalte verbalisieren beschreibt die Fähigkeit, Informationen und Sachverhalte situationsangemessen kommunizieren zu können	sich in der Gruppe austauschen können, Informationen, Planungen und Ergebnisse situationsangemessen, adressatengerecht, fachangemessen kommunizieren, chemische Fachsprache anwenden, Fachbegriffe kennen, Anwendung von Formelsprache	<i>Arbeite das bereitgestellte Informationsmaterial durch und berichte dann in eigenen Worten in deiner Gruppe darüber: Klopfestigkeit von Benzin – was steckt dahinter?</i>

	Erläuterung	Spezifizierung (ggf. aus chemischer Sicht)	Beispielaufgabe
Aufgabenanforderung	Umgang mit grafischen Repräsentationen beschreibt die Fähigkeit, relevante Informationen aus grafischen Darstellungen ziehen zu können bzw. Informationen und Daten grafisch darstellen zu können	grafische Darstellungen beschreiben und interpretieren, Daten in grafische Darstellungen umwandeln (pH-Kurven, Potentialkurven, Versuchszeichnungen, gezeichnete Darstellung bestimmter Sachverhalte)	<i>Ein Verfahren zur Ermittlung von Zusammensetzungen von Stoffgemischen ist die Gaschromatografie. Bei diesem Verfahren erhält man Diagramme (Gaschromatogramme), die man durch Vergleichen mit Diagrammen bekannter Stoffe deuten kann. Welche Aussagen kannst du anhand der Diagramme über die Zusammensetzung von Sommer- und Winterbenzin treffen?</i>
	Umgang mit Zahlen beschreibt die Fähigkeit, relevante Informationen aus numerischen Darstellungen ziehen zu können bzw. Informationen und Daten numerisch darstellen zu können	Rechnungen durchführen, die für die Chemie charakteristisch sind (Stöchiometrie, pH-Werte, Potentialberechnungen usw.), Größen(verhältnisse) abschätzen, bestimmte chemische Kenndaten in ihrer Größenordnung kennen, Informationsgehalt von Zahlen erkennen (Zahlentabellen anlegen und Informationen aus Zahlentabellen ziehen)	<i>n-Octan ist bei Zimmertemperatur eine Flüssigkeit mit der Dichte 0,703 g/ml. Gehe vereinfacht davon aus, dein Kraftstoff bestünde ausschließlich aus n-Octan und dein Fahrzeug benötige für 100 km Fahrt 6 Liter Treibstoff. Wie viel Kohlenstoffdioxid wird bei einer Fahrt von 100 km von deinem Motor ausgestoßen? Gib die Menge in Litern an!</i>
	Umgang mit mentalen Modellen beschreibt die Fähigkeit, Sachverhalte mit visuellen Vorstellungen verknüpfen zu können und die Grenzen dieser Vorstellung einschätzen zu können, Fähigkeit zur Verknüpfung von Sachverhalten mit räumlich-geometrischen Vorstellungen	Möglichkeit und Grenzen der Aussagekraft des Atommodells, Aufstellen von Reaktionsgleichungen, Verständnis für Strukturformeln, Teilchenmodell, Heranziehen von Modellvorstellungen zur Erklärung von Phänomenen wie Wärme- und elektrische Leitfähigkeit, ...	<i>Was passiert, wenn eine Flüssigkeit wie z. B. Wasser siedet? Stelle den Vorgang zeichnerisch da! Wie kannst du den dargestellten Vorgang auf molekularer Ebene erklären?</i>
	Bewerten beschreibt die Fähigkeit, sachliche Informationen in einen Bewertungsprozess einbinden zu können	Entscheidungen treffen und mit Hilfe der Chemie begründen, Informationsgehalt und -quellen bewerten, Bildung von Argumentationsketten (Pro-Kontra-Diskussionen), hierbei stellt die Chemie wiederum nur einen Schwerpunkt dar, in einen Bewertungsprozess kommen noch weitere Bewertungsgrundlagen zum Tragen (z. B. persönliche Vorlieben, ökonomische Aspekte usw.)	<i>Verschiedene Aspekte müssen bei der Einführung eines neuen Treibstoffs beachtet werden, wie z. B. Preis, Umweltverträglichkeit (Verbrennungsprodukte im Abgas), Verfügbarkeit, Gefährlichkeit usw. Diskutiere anhand dieser Aspekte, für welchen Treibstoff du dich entscheiden würdest!</i>

Tab. 6: verändert nach Deutsches PISA-Konsortium 2002; Aufgabenbeispiele entnommen und verändert nach Unterrichtsbeispielen, entwickelt in Zusammenarbeit mit den Schulsets in Hamburg und Rheinland-Pfalz

In Anlehnung an das internationale Konzept spielt bei der Erhebung 2003 die individuelle Wissensbasis des Aufgabenbearbeiters keine Rolle, da alle notwendigen Informationen im Aufgabentext zur Verfügung gestellt werden.

Abschließend wird im Zusammenhang mit der Kategorie Kompetenzanforderung ein Ansatz zur Aufgabensystematisierung von Müller & Müller (2002) vorgestellt. Sie unterscheiden verschiedene Aufgabentypen danach, welche Kompetenzen sie überprüfen bzw. fördern können, aber

auch sie weisen darauf hin, dass keine Aufgabe die beschriebene Funktion allein übernehmen kann, sondern dass der Erfolg, der mit einer Aufgabenstellung verbunden wird, in hohem Maße von der Vorbereitung durch und der Einbettung in den Unterricht abhängt (Müller et al. 2002). Nach ihrer Auffassung lassen sich vier Gruppen von Aufgabentypen ausmachen, die sich in Hinblick auf ihren Beitrag zur Kompetenzentwicklung, wie in Tabelle 7 aufgezeigt, untergliedern lassen.

Aufgabentyp	Beitrag zur Kompetenzentwicklung
Thematisierung von alltagsgebundenen Fehlvorstellungen und Erfahrung, Hilfe bei der Überwindung der Fehlvorstellungen	<ul style="list-style-type: none"> - anhand einer Alltagsbeschreibung können physikalische Fragestellungen erkannt werden; - physikalischer Zusammenhang eines Alltagsphänomens können beschrieben und erklärt werden; - Erklärungen können kommuniziert werden.
Aufbau von Routinen	<ul style="list-style-type: none"> - Lösungsalgorithmen und das Umstellen von mathematischen Gleichungen werden in geeigneter Weise beherrscht; - aus einem Text können physikalische Begriffe und Größen herausgelesen werden; - Einheiten werden richtig verwendet; - Messwerttabellen und Diagramme können aus physikalischer Sicht ausgewertet werden; - experimentelle Verfahren werden beherrscht.
Vernetzung und Festigung von Lerninhalten	<ul style="list-style-type: none"> - alte Kenntnisse können mit neuen Sachverhalten vernetzt werden; - kumulatives Lernen wird gefördert; - bekannte Prinzipien können auf neue Themenstellungen übertragen werden; - komplexe Aufgaben können durch elementare Modellvorstellungen bearbeitet werden.
Zulassen verschiedener Lösungswege, motivierende und sinnvolle Einbettung der Aufgaben in den Lernprozess	<ul style="list-style-type: none"> - mögliche Lösungswege können selbständig geplant werden; - Informationen können mithilfe unterschiedlicher Quellen und Medien beschafft werden; - das Bearbeiten der Aufgaben orientiert sich am naturwissenschaftlichen Denken und Arbeiten.

Tab. 7: Vier Aufgabentypen nach Müller et al. (2002) mit ihrem jeweiligen Beitrag zur Kompetenzentwicklung.

3.3.3.8 Zuordnung einer Aufgabe zu vorgegebenen Kompetenzstufen oder Kompetenzbereichen

Nach Vorschlag von Fischer et al. (2002) können Aufgaben insgesamt einer von sechs Kompetenzstufen zugeordnet werden, um auf diese Weise eine Aussage über die Schwierigkeit der Aufgaben zu treffen. Wiederum werden verschiedene mögliche Lösungswege getrennt betrachtet, da sich die Wege in ihrer Kompetenzstufe unterscheiden können. Diese Zuordnung eines Lösungsweges bzw. einer Aufgabe zu einer Kompetenzstufe oder einer Schwierigkeitsstufe kann abgesichert nur anhand statistischer Auswertungen erfolgen, anderenfalls muss diese Zuordnung als intuitiv oder theoriebasiert beschrieben werden.

Das internationale Konzept für die Erfassung naturwissenschaftlicher Kompetenz bei PISA beschreibt Prozesse, wobei unter Prozessen mentale und ggf. physische Handlungen verstanden werden, die beim Aufbau von Wissen und Verständnis ausgeführt werden. Im Rahmen des zu

Grunde gelegten Konzepts nach PISA werden fünf Aspekte von naturwissenschaftlicher Grundbildung betrachtet, die sich in den Aufgaben widerspiegeln.

Die Lernenden müssen zur Bearbeitung der Aufgaben naturwissenschaftliche Fragestellungen erkennen (1) und naturwissenschaftliche Nachweise identifizieren können (2), sie müssen die Fähigkeit besitzen, Schlussfolgerungen zu ziehen oder zu bewerten (3), gültige Schlussfolgerungen zu kommunizieren (4) und ein Verständnis von naturwissenschaftlichen Konzepten zeigen (5).

Eine „PISA-Aufgabe“ ist durch verschiedene charakteristische Merkmale gekennzeichnet: Typisch ist ihr verhältnismäßig umfangreicher Stimulus in Form von Texten und/oder Abbildungen. Da das Faktenwissen bei der Erfassung von naturwissenschaftlicher Bildung keine Rolle spielen soll, müssen alle nötigen Informationen, die zum Lösen einer Aufgabe notwendig sind, mitgeliefert werden. Um den Umfang des „Stimulus“ zu rechtfertigen, enthält eine Aufgabe immer mehrere Teilaufgaben.

Die *Situation* einer veröffentlichten Aufgabe umfasst z. B. das Thema des Klonens, damit stellt eine aktuelle, gesellschaftliche Diskussion die Oberflächenstruktur der Aufgabe dar. Um die Anbindung an die Lebenswelt noch augenscheinlicher zu gestalten, ist der Aufgabentext als Zeitungsartikel verfasst.

Um die einzelnen Teilaufgaben zu lösen, müssen verschiedene naturwissenschaftliche *Prozesse* ausgeführt werden, und generell erfordert die Bearbeitung einer Aufgabe Kenntnisse zu grundlegenden *Konzepten* (Zusammenstellung der Konzepte s. Deutsches PISA-Konsortium 2002).

Beide Ansätze der nationalen und internationalen PISA-Erhebung sind darauf ausgelegt, kognitive Kompetenzen bzw. Prozesse zu erfassen, die sich auf die Naturwissenschaften insgesamt beziehen. Die Chemie kommt in den Aufgaben nicht in gleichem Maße zum Tragen wie die Biologie und die Physik, da der Chemieunterricht in der Regel an deutschen Schulen bisher erst in der neunten Klasse beginnt. Dennoch lassen sich diese Aufgabenanforderungen auf die Chemie anwenden, wenn das Charakteristische der Chemie, insbesondere bei den kognitiven Kompetenzen, berücksichtigt wird (vgl. Tab. 6, Kap. 3.3.3.7).

Es wurde bereits darauf verwiesen, dass es insbesondere in Hinblick auf Vergleichsarbeiten, Standards und Zentralabitur wichtig ist, mit grundlegenden, transparenten und vor allem verbindlichen Kriterien bei der Konstruktion und Analyse von Aufgaben zu arbeiten. In den nationalen Bildungsstandards für das Fach Chemie (KMK 2004a) sind vier Kompetenzbereiche (Fachwissen, Erkenntnisgewinnung, Kommunikation und Bewertung) formuliert worden, die sich in ihren Anforderungen in drei Niveaus untergliedern lassen.

Die Zuordnung einzelner Aufgaben zu diesen drei Anforderungsbereichen erfolgt wiederum anhand einer theoriebasierten Einschätzung, da bisher keine Kompetenzmodelle zur Verfügung stehen, die eine Einordnung auf einer empirisch gesicherten Grundlage ermöglichen. Mit Hilfe eines solchen vorgegebenen, inhaltlich ausgearbeiteten Kriterienrasters werden Aufgabenbeispiele entwickelt und klassifiziert (dazu genauer Kap. 4.2).

3.3.3.9 Sonderfall: Beispielaufgaben

Eine Sonderstellung bei der Betrachtung von Aufgaben nimmt der Aufgabentyp „Beispielaufgabe“ (auch „ausgearbeitete Lösungsbeispiele“) ein, die sich aufgrund ihrer grundlegend anderen Zielstellung nicht sinnvoll in die vorgestellten Ansätze zur Aufgabenanalyse und -konstruktion einordnen lassen. Beispielaufgaben zeichnen sich dadurch aus, dass sie neben der eigentlichen Aufgabenstellung eine mehr oder weniger detaillierte Darstellung des Lösungsweges enthalten, d. h. sie beschreiben die Anwendung der zur Lösung notwendigen Konzepte (Renkl 2001, Mackensen & Sandmann 2002).

Auf diese Weise lernen beim Durcharbeiten von ausgearbeiteten Lösungsbeispielen insbesondere Schülerinnen und Schüler mit einem geringen fachspezifischen Wissen effektiver und mit größerem Lernerfolg als durch selbständiges Problemlösen, denn sie empfinden in geringerem Maße Überforderung und Frustration, und sie erfahren in höherem Maße eigenes Kompetenzerleben (Mackensen et al. 2002). In engem Zusammenhang mit Beispielaufgaben wird die Methode des Selbsterklärens gesehen, denn die Art, wie sich die Lernenden mit einer Beispielaufgabe auseinandersetzen (z. B. Durchlesen der Beispielaufgabe, Schritte in eigenen Worten paraphrasieren, Bezüge herstellen oder Lösungsschritte antizipieren), ist maßgeblich für den Erfolg des Lernens mit Lösungsbeispielen.

Insgesamt wird auch hier deutlich, dass das Vorwissen der Schülerinnen und Schüler berücksichtigt werden muss. Mackensen et al. (2002) schlagen vor, die Lernimpulse in zwei Versionen zu formulieren, um Beispielaufgaben für Lernende mit geringem und mit hohem fachspezifischem Wissen bereitstellen zu können. Diese Form von Aufgaben zeichnet sich dadurch aus, dass sie eine relativ einfache Differenzierung und damit eine Anpassung an den Leistungsstand der Lernenden ermöglicht, und eignet sich besonders zur Förderung der Problemlösefähigkeit und zum Erlernen neuer und Vertiefung bereits erarbeiteter Sachverhalte.

3.3.3.10 Zusammenfassung

Aufgaben werden in einer weit gefassten Definition als „Aufforderungen zum Lernhandeln“ verstanden, die sich dem Lernenden in drei Komponenten Ausgangssituation, Transformation und Endsituation präsentieren. Die Aufgabenkultur beschreibt das Gesamtgefüge von Unterrichtsinhalt, Unterrichtsmethode, Aufgabengestaltung und Intention der Aufgabe.

Wird der von Fischer et al. (2001, 2002) vorgeschlagene Ansatz verfolgt, ergeben sich für einzelne Aufgaben so genannte „Datenblätter“, die ausführliche Informationen über die einzelne Aufgabe liefern. Darüber hinaus besteht die Möglichkeit, anhand dieser „Datenblätter“ Aufgaben miteinander zu vergleichen oder Ansatzpunkte für Veränderungen zu identifizieren.

Für die Analyse einer gesamten Aufgabenkultur ist dieser sehr umfassende Ansatz vor allem für den Schulalltag zu aufwendig. Zudem ist es wichtig, sich vor der Analyse von Aufgaben das Ziel der Analyse und damit die Zweckmäßigkeit der ausgewählten Kategorien, nach denen die Aufgaben untersucht werden, bewusst zu machen. Nachdem das Ziel der Untersuchung deutlich gemacht ist, muss die Frage nach der Zweckmäßigkeit des verfolgten Systematisierungsansatzes beantwortet werden. Es ist nicht notwendig und auch nicht sinnvoll, in einem Aufgabenfundus nach verschiedenen Aufgabentypen zu „fahnden“, wenn es vielmehr darum geht, zu untersuchen, ob Aufgaben auf vielfältige Weise (und nicht immer nur in Übungsphasen als Stillarbeit oder Hausaufgaben) in den Unterricht eingebunden sind.

Während die Beachtung von eindeutigen Operatoren und präzisen Formulierungen in einer Aufgabe mit dem Ziel der Leistungsbewertung von großer Bedeutung ist, kann es für eine Aufgabe, die den Einstieg in ein Thema verfolgt und in erster Linie Interesse wecken und für die Facetten des Themengebietes sensibilisieren soll, zweckmäßig sein, gerade auf präzise Formulierungen im Aufgabentext zu verzichten.

Zusammenfassend ist festzuhalten, dass es einer Auswahl zweckmäßiger Kriterien bedarf, um eine Klassifizierung von Aufgaben und ihres Einsatzes im Unterricht zu ermöglichen. Diese Kriterien müssen sich aus der Zielstellung des Unterrichts oder der Untersuchung ergeben. Die Ziele eines Unterrichts, der sich an der Konzeption *Chemie im Kontext* orientiert, sind bereits in Kap. 2.1 dargestellt worden. Anhand dieser Ziele werden in den folgenden Abschnitten dieser Arbeit Kriterien abgeleitet, die für eine Entwicklung und Untersuchung von Aufgaben, die für diese Form des Unterrichts typisch und geeignet sind, Anwendung finden können.

Es bleibt insgesamt festzuhalten, dass bei einer ausführlichen Betrachtung von Aufgaben und Aufgabenkultur alle genannten Bereiche berücksichtigt werden sollten, weil nur die Gesamtheit ein vollständiges Bild liefert.

3.4 Was ist eine „Chemie-im-Kontext-Aufgabe“? Versuch einer Definition

Aus dem Vorgegangenen wird deutlich, dass es wenig sinnvoll ist, eine Aufgabe isoliert zu betrachten, um sie dann z. B. einer bestimmten Form des Unterrichtens zuzuordnen. Sicherlich gibt es Aufgaben, die sich stärker durch einen „Chemie-im-Kontext-Charakter“ auszeichnen als andere, aber um sie letztlich als „Chemie-im-Kontext-Aufgabe“ bezeichnen zu können, ist es notwendig, den gesamten Unterricht mit Inhalten und Methoden zu betrachten. Denn nur mit dem tatsächlichen Unterricht, in den eine Aufgabe eingebettet ist, erhält sie ihren „Chemie-im-Kontext-Charakter“.

Verschiedene Merkmale, die sich wiederum den einzelnen oben vorgestellten Kategorisierungen zuordnen lassen, machen eine Aufgaben zu einer „Chemie-im-Kontext-Aufgabe“. In der Tabelle 8 wird verdeutlicht, dass sich die Beschreibungen in einzelnen Kategorien auf naturwissenschaftlichen Unterricht insgesamt anwenden lassen und natürlich nicht speziell für einen Unterricht nach *Chemie im Kontext* gelten (wie z. B. Art der möglichen Lösungswege, Antwortformate und das Experimentierverhalten).

Kategorie	bedeutsam u. a. für eine „Chemie-im-Kontext-Aufgabe
Inhalt, Lehrplan oder Richtlinien	<ul style="list-style-type: none"> - der Kontext besitzt Bedeutung in der inhaltlichen Aufgabengestaltung, - die Aufgabe erschließt einzelne Aspekte eines Kontextes, - der Kontext erweist sich in der Aufgabe als tragfähig (dient nicht nur als „Aufhänger“), - die Aufgabe fließt in die gesamte Kontexterarbeitung ein und steht nicht isoliert, - die Aufgabe verbindet kontextliche und fachliche Aspekte, - die Aufgaben stellen einen engen Bezug zu Basiskonzepten her.
Unterrichtsphasen	<ul style="list-style-type: none"> - die Aufgabe ist konstitutiv für eine Phase des Unterrichts (Begegnung, Planung, Durchführung, Vertiefung und Vernetzung) bzw. lässt sich einer der Unterrichtsphasen im Rahmen einer Kontexterarbeitung zuordnen, - in der Phase der Leistungsüberprüfung wird sie allen Aspekten des vorangegangenen Unterrichts gerecht, - durch die Aufgabe wird ausreichend Möglichkeit zur Übung und Festigung neuer Inhalte gegeben.
Art der möglichen Lösungswege	<ul style="list-style-type: none"> - die Aufgabe lässt verschiedene Lösungswege zu und forciert nicht bereits durch ihre Formulierungen einen „richtigen“ Weg, - die Lehrkraft hat nicht bereits eine „fertige Lösung“ im Kopf.
Antwortformate	<ul style="list-style-type: none"> - in verschiedenen Aufgaben werden unterschiedliche Formate berücksichtigt, - bei der Auswahl des Formats wird die Zweckmäßigkeit beachtet, - in der Aufgabengesamtheit wird eine Vielfalt in den Antwortformaten angestrebt.
Grad der Offenheit	<ul style="list-style-type: none"> - in Aufgaben werden durch Art der Lösungswegvorgabe Möglichkeiten zur Differenzierung ausgenutzt, d. h. der Spielraum zwischen Vorgabe eines bestimmten Lösungsweges und Freistellen des Lösungsweges wird ausgeschöpft, - Lernende werden bereits bei der Aufgabengestaltung einbezogen.
Experimentierverhalten	<ul style="list-style-type: none"> - zweckgebunden wird durch die Aufgabenformulierung imitatorisches (z. B. Einüben von bestimmten Experimentierfertigkeiten), organisierendes (z. B. Arbeit in Gruppen üben) und konzeptuelles Experimentieren (naturwissenschaftliches Arbeiten) vorgegeben, - Ziel ist die Förderung eines konzeptuellen Experimentierverhaltens, - die Experimente zeichnen sich durch einen tragenden Kontextbezug aus.
Anforderungsmerkmale	<ul style="list-style-type: none"> - neben dem Fachwissen werden verschiedene Kompetenzen, die zentral für naturwissenschaftliches Denken und Arbeiten sind, berücksichtigen (dazu genauer Kap. 3.3.3.7, Tab. 6).
Kompetenzstufen oder Kompetenzbereiche	<ul style="list-style-type: none"> - bei der Aufgabenformulierung werden verschiedene zentrale Kompetenzbereiche berücksichtigt, - Aufgaben lassen sich verschiedenen Kompetenzbereichen und auch Kompetenzstufen zuordnen.

Tab. 8: Bedeutsame Aspekte einer „Chemie-im-Kontext-Aufgabe“, untergliedert nach den in den Kap. 3.3.3.1 – 3.3.3.8 beschriebenen Kategorien.

Die hier vorgenommene Auflistung zeigt, welche Anforderungen anhand der Zuordnungskategorien an eine Aufgabe gestellt werden, die sich als typisch und geeignet für einen Unterricht nach *Chemie im Kontext* charakterisieren lässt. Zugleich wird aber auch deutlich, dass es Anforderungen sind, die insgesamt an eine „gute“ Aufgabe und eine „gute“ Aufgabenkultur gestellt werden (Kap. 3.3.2). Eine spezifische „*Chemie-im-Kontext-Aufgabe*“ wird eine Aufgabe maßgeblich durch Erfüllung der ersten beiden Aspekte Inhalt und Unterrichtsphase und damit durch den tatsächlichen Unterricht, in den sie eingebunden ist.

3.5 Kompetenzerfassung – Hinweise zur Testkonstruktion und Aufgabenentwicklung

Aus verschiedenen Arbeiten zur Wissens- und Kompetenzerfassung im naturwissenschaftlichen Unterricht werden im Folgenden Aspekte zusammengestellt, die bei der Konstruktion eines Testinstruments und der Entwicklung von Aufgaben zur Erfassung von Kompetenzen berücksichtigt werden müssen. Dabei wird neben allgemeinen Hinweisen (Häußler, Bündler, Duit, Gräber & Mayer 1998, Klauer 2001, Duit et al. 2001,) wiederum auf bereits erwähnte Arbeiten eingegangen, die sich auf internationale und nationale PISA-Erhebungen und Ausführungen zur Einführung von nationalen Bildungsstandards (BMBF 2007) beziehen.

Mit einer Leistungsüberprüfung auf Seiten der Lernenden geht immer zugleich eine Bewertung des Unterrichtserfolges einher, da auf diese Weise das Erreichen der gesetzten Ziele überprüft wird. In diesem Zusammenhang wird erneut deutlich, dass es von zentraler Bedeutung ist, dass alle Unterrichtsziele in die Bewertung der Lernenden und damit auch des Unterrichts einbezogen werden (Häußler et al. 1998b). Für Leistungskontrollen gelten die Gütekriterien, die an jedes Messinstrument angelegt werden: Objektivität, Reliabilität und Validität. Insgesamt lassen sich in Hinblick auf diese Gütekriterien folgende Aspekte zusammenfassen (Häußler et al. 1998b): Multiple-Choice-Aufgaben zeichnen sich durch eine hohe Objektivität aus, während Aufgaben mit offenen Antwortformaten bei der Bewertung der Antwort einen bedeutend größeren Spielraum ermöglichen. Dieser Spielraum lässt sich einschränken, indem der Erwartungshorizont und damit die Auswertungskriterien sehr detailliert beschrieben werden. Ein wichtiges Kriterium innerhalb der Reliabilität eines Tests ist die Trennschärfe von Aufgaben. Aufgaben, die zu leicht sind, werden von allen Probanden gelöst, während zu schwierige Aufgaben von niemandem korrekt gelöst werden. Anhand solcher Aufgaben lassen sich keine Aussagen zum tatsächlichen Leistungsstand treffen, sie sind nicht trennscharf. Aussagen zur Validität eines Tests sind schwierig zu treffen, vor allem, wenn die Lernwege von Probanden nicht genau bekannt sind. Es lässt sich nicht mit Sicherheit sagen, ob Antworten zu einzelnen Aufgaben ein Reproduzieren von Gelerntem oder eine aktive Bearbeitung widerspiegeln.

Ein Expertenrating kann als Mittel zur Validitätsschätzung herangezogen werden, indem untersucht wird, ob sich in dem Testinstrument theoretisch belegte Elemente wiederfinden, die als Indikatoren für bestimmte Fähigkeiten, die gemessen werden sollen, geeignet sind. Eine zweite Möglichkeit zur Schätzung der Validität eines Tests ist das Verfahren der Faktorenanalyse, denn durch sie lässt sich ermitteln, ob ein Test ein- oder mehrdimensional misst.

Häußler et al. (1998b) und Graf (2001) haben zusammengestellt, inwieweit geschlossene, halboffene und offene Antwortformate sich zur Überprüfung bestimmter kognitiver Fähigkeiten eignen. Zusammenfassend wird festgehalten, dass sich Aufgaben mit einem geschlossenen Antwortformat dazu eignen, Wissen über Einzelheiten, Benennungen, über Begriffe und Aussagen zu erfassen. Aufgaben mit einem halboffenen Antwortformat eignen sich dazu, Wissen über Einzelheiten, Benennungen, über Begriffe, Aussagen und das Verstehen von Zusammenhängen und Prozessen zu erfassen, während Aufgaben mit einem offenen Antwortformat dazu geeignet sind, das Verstehen von Zusammenhängen und Prozessen, höhere kognitive Leistungen und die Fähigkeit zum Bewerten, Abwägen und Argumentieren zu erfassen.

Sie geben darüber hinaus den Hinweis, dass es günstig ist, einen Test mit einer so genannten „Eisbrecher“-Aufgabe (einer sehr leichten Aufgabe) zu beginnen, um allen Lernenden zunächst ein Erfolgserlebnis zu ermöglichen (Häußler et al. 1998b, dort auch weitere Hinweise zu Besonderheiten der einzelnen Aufgabentypen).

Aufgaben mit offenem Antwortformat haben den Vorteil, dass sich z. B. Denkfehler im Lösungsweg besser rekonstruieren lassen. Nachteile offener Aufgaben sind der hohe Aufwand und die verminderte Objektivität der Auswertung. Bei geschlossenen Aufgaben stellt die Auswertobjektivität und auch die technische Erfassung kein Problem dar, Nachteile sind die begrenzte Eignung zur Erhebung komplexerer kognitiver Fähigkeiten und die Tatsache, dass die korrekte Antwort zu einem gewissen Prozentsatz vom Probanden geraten werden kann (Klauer 2001). Zudem stellt das Formulieren geeigneter und anspruchsvoller Distraktoren eine große Herausforderung für den Aufgabenentwickler dar, vor allem dann, wenn über Multiple-Choice-Aufgaben z. B. Fehlvorstellungen aufgespürt werden oder wenn Distraktoren Aufschluss über den gewählten Lösungsweg geben sollen (Klauer 2001, Duit et al. 2001).

Im Rahmen der PISA-Erhebungen wird u. a. aus Gründen der Handhabung überwiegend mit geschlossenen Antwortformaten gearbeitet (Kap. 3.3.3.8). In der ersten nationalen Zusatzerhebung 2000 hat sich gezeigt, dass deutsche Schülerinnen und Schüler mit Aufgaben gut zurecht kommen, wenn aus grafisch präsentierten Aufgabeninhalten Informationen abgeleitet werden müssen, besondere Schwierigkeiten treten auf, wenn Faktenwissen auf neue, unbekannte Situa-

tionen übertragen und angewendet oder mentale Modelle zur Bearbeitung einer Aufgabe herangezogen werden müssen (Prenzel et al. 2002).

Diese und weitere Ergebnisse zu Aufgabenelementen und Anforderungen, die den Schwierigkeitsgrad einer Aufgabe beeinflussen (Prenzel et al. 2002), sind in der folgenden Tabelle 9 zusammengestellt, wobei die Autoren aber vor einer Generalisierung der Ergebnisse warnen, denn eine grafische Darstellung kann die Aufgabenschwierigkeit natürlich auch deutlich erhöhen.

Aufgabenschwierigkeit	Element der Aufgabe, Aufgabenanforderung
Deutschen Schülerinnen und Schüler fällt es verhältnismäßig leicht, Aufgaben zu lösen,...	... in denen aus graphischen Darstellungen Informationen abgeleitet werden müssen, weil es die Anforderungen an das Verstehen des Textes senkt.
Keinen Einfluss auf die Schwierigkeit einer Aufgabe hat...	... die Länge des Aufgabentextes. ... es, wenn die Schülerinnen und Schüler konvergent denken, d. h. wenn Informationen logisch miteinander verknüpft werden müssen.
(Große) Schwierigkeiten bereiten den deutschen Schülerinnen und Schülern Aufgaben, in denen Faktenwissen auf neue Situationen übertragen und angewendet werden muss. ... in denen mentale Modelle zur Bearbeitung herangezogen werden müssen. ... in denen zur Bearbeitung aktiv formuliert werden muss. ... in denen etwas ausgerechnet werden muss. ... in denen divergentes Denken gefordert ist. ... in denen sie sich etwas räumlich vorstellen müssen. ... in denen das Nennen oder Verstehen eines bestimmten naturwissenschaftlichen Begriffes der erfolgreiche Auseinandersetzung zu Grunde gelegt wird.

Tab. 9: Einfluss einzelner Aufgabenelemente auf die Schwierigkeit einer Aufgabe (Prenzel et al. 2002).

Erhebungen im Rahmen der internationalen PISA-Untersuchungen haben gezeigt, dass die eingesetzten Testaufgaben naturwissenschaftliche Grundbildung anhand einer Skala abbilden, die das gesamte Spektrum wiedergibt, d. h. von dem Erkennen von Fragestellungen, zu deren Beantwortung naturwissenschaftliche Untersuchungen einen Beitrag leisten können, bis hin zu einem Verständnis und zur Anwendung abstrakter und komplexer naturwissenschaftlicher Konzepte, um Sachverhalte zu erklären. Bei der Erstellung von Tests zur Kompetenzerfassung muss berücksichtigt werden, dass die einzelnen als naturwissenschaftlich angesehenen Fähig- und Fertigkeiten nicht ohne Weiteres unabhängig voneinander untersucht werden können. Auch im Rahmen der ersten PISA-Studie wird zunächst von nur einer Skala des naturwissenschaftlichen Verständnisses ausgegangen, wobei berücksichtigt werden muss, dass die Naturwissenschaften in diesem Untersuchungszeitraum keinen Schwerpunkt darstellten. Darüber hinaus werden Aufgabenzuordnungen zu einzelnen Niveaustufen naturwissenschaftlicher Bildung nicht vorab vorgenommen, sondern anhand empirischer Daten im Nachhinein. Es ist folglich problematisch, Aufgaben theorie- oder erfahrungsbasiert bestimmten Schwierigkeitsstufen zuzuordnen, solange keine empirisch abgesicherten Kompetenzmodelle vorliegen.

In der Expertise zur Einführung und Entwicklung von nationalen Bildungsstandards (BMBF 2007) werden allgemeine Empfehlungen gegeben, indem sie deutlich darauf hinweisen, dass die

Kompetenzen, die erfasst werden sollen, zunächst so genau und konkret beschrieben werden müssen, dass sie sich in Aufgabenstellungen übertragen lassen und durch Testverfahren erfasst werden können. Dazu ist es erforderlich, empirisch gesicherte Kompetenzmodelle zu entwickeln, anhand derer dann mit Hilfe passender Aufgaben das Erreichen einer bestimmten Kompetenzstufe überprüft werden kann. Diese passenden Aufgaben müssen demnach valide Indikatoren für das Erreichen einer Kompetenzstufe darstellen, d. h. das Anforderungsprofil einer Aufgabe muss das erfassen, was für eine bestimmte Kompetenzstufe charakteristisch ist.

Zusammenfassend mit Blick auf die im Rahmen der vorliegenden Arbeit zu entwickelnden Testaufgaben lässt sich festhalten,

- dass der Test nicht ausschließlich aus Aufgaben mit offenen Antwortformaten bestehen sollte, sondern dass auch verschiedene Formen des geschlossenen Antwortformats berücksichtigt werden.
- Die Antworten bei offenen Formaten sollten nicht nur aus Textantworten bestehen, sondern auch andere Formen wie z. B. Zeichnungen und Rechnungen berücksichtigen.
- Darüber hinaus wird neben bestimmten Fähigkeiten auch Wissen zur Aufgabenbearbeitung erfordert, d. h. anders als bei PISA-Aufgaben werden nicht alle notwendigen Informationen im Aufgabentext bereitgestellt.
- Für die Antworten der offenen Antwortformate wird ein detaillierter Erwartungshorizont formuliert, um ein möglichst hohes Maß an Auswertungsobjektivität gewährleisten zu können.
- Um Aussagen darüber treffen zu können, welche Aufgabeninhalte den Schülerinnen und Schülern aus dem vorangegangenen Unterricht bekannt sind, werden die unterrichtenden Lehrkräfte befragt (dazu genauer Kap. 6.3.2).
- Auf eine Zuordnung der einzelnen Aufgaben bzw. Lösungswege zu Kompetenzniveaus wird verzichtet, die Schwierigkeiten der einzelnen Aufgabe werden im Anschluss an die Untersuchung statistisch bestimmt.

4 Entwicklung von Kategorien zur Analyse und Konstruktion von Aufgaben, die den Anforderungen eines Unterrichts nach *Chemie im Kontext* gerecht werden

Im folgenden Kapitel werden die Überlegungen zur Entwicklung von Kategorien zur Konstruktion und Analyse von Aufgaben ausgeführt, die sich inhaltlich aus den Zielen eines Unterrichts nach *Chemie im Kontext* herleiten lassen. Die Ziele der Unterrichtskonzeption sind in Kapitel 2.1 bereits dargestellt worden und werden an dieser Stelle in kurzer Form erneut aufgegriffen. Zunächst werden die inhaltlichen Kriterien konkretisiert und anhand von Beispielen veranschaulicht, die sich aus den Grundlagen des naturwissenschaftlichen Denkens und Arbeitens und aus den kognitiven Aufgabenanforderungen, die im Rahmen der nationalen PISA-Erhebung 2003 realisiert wurden, ergeben. Im Anschluss daran werden diese beiden Dimensionen „Erkenntnis-schritte“ und „kognitive Anforderungen“ zu einem Modell zusammengeführt und die Handhabung dieses Modells diskutiert.

Abschließend wird anhand der theoretischen Grundlagen die eingesetzten Testaufgaben vorgestellt und diskutiert. Die einzelnen Testitems werden in Hinblick auf ihre Gestaltung analysiert und das Testinstrument in seiner Gesamtheit beschrieben.

4.1 Inhaltliche Kriterien zur Konstruktion und Analyse von Aufgaben

In Kapitel 2.1 wurde aufgezeigt, welche Ziele die Konzeption *Chemie im Kontext* verfolgt. Aus der Orientierung an verschiedenen Aspekten von naturwissenschaftlicher Grundbildung (u. a. Kompetenzbereiche der Bildungsstandards, „Nature of Science“) werden Kompetenzbereiche und kognitive Aufgabenanforderungen zusammengefasst und abgeleitet, durch die sich die Zielstellung der Konzeption in Aufgaben konkreter beschreiben lässt. Die auf diese Weise gewonnenen kognitiven Aufgabenanforderungen und Kompetenzbereiche stellen den Rahmen für die Entwicklung eines Kategorienmodells zur Analyse und Konstruktion von Aufgaben dar, die insbesondere geeignet sind, das Erreichen dieser Ziele, die zentral in einem Unterricht nach *Chemie im Kontext* sind, transparent und überprüfbar zu machen.

4.1.1 Erste Dimension:

Naturwissenschaftliches Denken und Arbeiten in *Chemie im Kontext*

Die folgenden Ausführungen beziehen sich auf die Prozessschritte, die bereits in Kapitel 2.2 eingeführt wurden (siehe Abb. 2, S. 21). Sie stellen eine zentrale Dimension bei der Analyse und Konstruktion von Aufgaben dar und werden in diesem Abschnitt beispielhaft genauer erläutert.

Diese hier dargestellten Schritte beschreiben Kompetenzen, die durch einen Unterricht nach *Chemie im Kontext* gefordert und gefördert werden. Sie zeigen Facetten des naturwissenschaftlichen Arbeitens auf, die entsprechend der Zielstellung der Konzeption an die Schülerinnen und Schüler vermittelt werden. Um den Lernstand einer Klasse hinsichtlich des Erreichens dieses Ziels überprüfen zu können, bedarf es Aufgaben, die diese Aspekte in ihren Aufgabenanforderungen widerspiegeln. Diese Schritte liefern demnach Kategorien, nach denen die Aufgaben für das Testinstrument entwickelt werden.

Die folgende Tabelle 10 konkretisiert die beschriebenen Schritte anhand von Beispielen.

Arbeitsschritte in einem Unterricht nach <i>Chemie im Kontext</i>	Beispiel aus einer Unterrichtseinheit zum Thema <i>Treibstoffe – Benzin</i>
Kontext erschließen, Fragen zum Kontext formulieren	Warum ist gerade Benzin als Treibstoff so geeignet? Vor- und Nachteile, Alternativen, Sicherheitsaspekte, politische, ökologische, ökonomische Aspekte...
Fragen identifizieren, die chemisch untersucht werden können	Was passiert im Motor? Was ist Benzin? Welchen Anforderungen muss Benzin genügen?...
Vorgehensweise überlegen und ggf. Experimente oder andere Untersuchungen planen	Planung von Experimenten (Flüchtigkeit, Siedebereiche, Zusammensetzung), Internet- und Literaturrecherchen (Funktionsweise des Motors...)
Experimente oder andere Untersuchungen durchführen	Durchführung der geplanten Untersuchungen zu Eigenschaften des Benzins (Analyse der Verbrennungsprodukte, Untersuchung der Löslichkeit, Siedepunktsbestimmung usw.)
Ergebnisse darstellen und kommunizieren, ggf. Hypothesen überprüfen	Zusammenführung der Ergebnisse und Ableitung von Kriterien zur Beurteilung eines Treibstoffs: Brennbarkeit/ Verbrennungsprodukte, Verfügbarkeit, Optimierbarkeit, Siedepunkte, ...
Ergebnisse bewerten und in den Kontext einordnen	anhand der ermittelten Kriterien können zur Überprüfung und erneuten Anwendung des erarbeiteten Wissens Treibstoffalternativen vergleichend beurteilt werden
Ergebnisse bewerten und in fachchemischen Zusammenhang einordnen	unterschiedliche Siedepunkte von Stoffen können anhand der unterschiedlichen Molekülstrukturen der Stoffe und der daraus ableitbaren intermolekularen Wechselwirkungen erklärt werden
Einzelergebnisse zu übergeordneten Prinzipien abstrahieren	Abstraktion: durch die Struktur von Stoffen lassen sich Eigenschaften und typische Reaktionen von Stoffen erklären und vorhersagen

Tab. 10: Konkretisierung der Prozessschritte anhand von Beispielen aus einer Unterrichtseinheit zum Kontext „Benzin“. Diese Einheit wurde im Rahmen der Projektarbeit entwickelt (Kap. 2.3.2, Nentwig et al. 2004).

4.1.2 Zweite Dimension:

Kognitive Aufgabenanforderungen im Bereich der Naturwissenschaften

Die zweite Dimension des im Rahmen dieser Arbeit ausgearbeiteten Modells zur Analyse und Konstruktion der Aufgaben stellen konkrete kognitive Aufgabenanforderungen dar, die für naturwissenschaftliches Denken und Arbeiten zentral sind. Diese Anforderungen werden im folgenden Abschnitt genauer erläutert.

Im Rahmen der nationalen PISA-Zusatzerhebung 2003 sind für die Aufgabenkonstruktion kognitive Anforderungen formuliert worden, die zusammen mit den „Big ideas“ der Fächer Biologie, Chemie und Physik den Rahmen für die eingesetzten Aufgaben darstellen (Kap. 3.3.3.7 der vorliegenden Arbeit). An diesen Anforderungen orientiert sich auch die Konstruktion der in dieser Arbeit entwickelten Testaufgaben. Die kognitiven Anforderungen wurden zu diesem Zweck in einigen Fällen „chemisch“ ausdifferenziert (Kap. 3.3.3.7, Tab. 6), denn im Rahmen der nationalen PISA-Erhebung 2003 ist das nicht geschehen, da sich die Chemiekenntnisse von 15-Jährigen (zumeist Neuntklässern) auf wenige Kenntnisse aus dem Chemieanfangsunterricht beschränken.

4.2 Zusammenführung zu einem Modell zur Beschreibung von Aufgaben

Die in den vorangegangenen Abschnitten vorgestellten Erkenntnisschritte in einem Unterrichtsgang nach *Chemie im Kontext* und die sieben kognitiven Aufgabenanforderungen bilden die zwei Dimensionen des erarbeiteten Modells für die Konstruktion von Aufgaben, das im Folgenden der Auseinandersetzung mit Aufgaben zu Grunde gelegt wird und sich nach den Zielen der Unterrichtskonzeption *Chemie im Kontext* richtet. Im folgenden Abschnitt werden beide Dimensionen zusammengeführt, als Gesamtmodell vorgestellt und mit Beispielen erläutert.

An diesem Modell orientiert sich die Formulierung der Testaufgaben für das zentrale Untersuchungsinstrument (Kap. 4.3), aber neben dem in dieser Arbeit überwiegenden Konstruktionsaspekt bietet sich dieses Instrument auch zur Analyse von bestehenden Aufgaben an. Eine solche Analyse kann beispielsweise Anhaltspunkte dazu liefern, welche Kompetenzbereiche und Aufgabenanforderungen in einer Stationsarbeit oder aber auch in einer Leistungsüberprüfung zu großes Gewicht besitzen oder vernachlässigt werden.

Tabelle 11 zeigt die Zusammenführung der beiden Dimensionen.

	divergentes Denken	konvergentes Denken	Umgang mit grafischen Darstellungen	Umgang mit Sprache	Umgang mit Zahlen	Bewerten	Umgang mit mentalen Modellen
Kontext erschließen, Fragen formulieren							
Fragen identifizieren, die chemisch untersucht werden können							
Vorgehensweise planen							
Untersuchungen durchführen							
Ergebnisse darstellen und kommunizieren							
Ergebnisse bewerten und einordnen (Kontext)							
Ergebnisse bewerten und einordnen (Chemie)							
Abstraktion zu Basis-konzepten							

Tab. 11: Zusammenführung der Dimensionen „Prozessschritte“ und „kognitive Aufgabenanforderungen“.

Theoretisch lassen sich für alle „Felder“ der Tabelle Aufgaben formulieren, aber nicht alle Kombinationen sind gleich sinnvoll. Zudem lassen sich die Felder nicht immer eindeutig voneinander trennen, vielmehr handelt es sich um fließende Übergänge mit festgelegten Dominanzen bestimmter Aufgabenaspekte, so dass sich z. B. durch geringfügige Veränderungen in der Aufgabengestaltung andere Schwerpunkte setzen lassen.

Die Tatsache, dass sich Aufgaben z. T. zahlreichen Feldern dieser Tabelle zuordnen lassen, zeigt, mit welcher Komplexität in den Anforderungen die Schülerinnen und Schüler in der jeweiligen Aufgabe konfrontiert werden. In diesem Zusammenhang kann es hilfreich sein, die von Fischer et al. (2002) vorgeschlagene und in den Kapiteln 3.3.3.1 – 3.3.3.8 vorgestellte Zuordnung vorzunehmen, die für jede einzelne Aufgabenanforderung entscheidet, in welchem Maße sie für die Bearbeitung der Aufgabe von Bedeutung ist.

Während einerseits einzelne Kombinationen in der Tabelle weniger sinnvoll erscheinen als andere (wird z. B. bereits in der Erschließungs- und Planungsphase ein Schwerpunkt auf den Umgang mit Modellen gelegt, wird man der geforderten Offenheit für diese Phase nicht gerecht), zeichnen sich andererseits bestimmte Prozessschritte durch einen Schwerpunkt einer kognitiven Anforderung aus (z. B. „divergentes Denken“ bei „Kontext erschließen, Kontextfragen formulieren“, „konvergentes Denken“ bei „Fragen zur Bearbeitung formulieren“).

Wie eine Analyse von Aufgaben aus Chemie-Schulbüchern verdeutlicht, lassen sich einige Felder verhältnismäßig leicht mit Aufgabenbeispielen füllen, da sie bereits im herkömmlichen Chemieunterricht einen großen Stellenwert besitzen (z. B. konvergentes Denken und Umgang mit Modellvorstellungen bei der Bewertung und Einordnung von gewonnenen Ergebnissen in

den Bereich der Chemie). Das Formulieren von Aufgaben für andere Tabellenfelder stellt höhere Anforderungen an den Aufgabensteller, weil sie für den herkömmlichen Unterricht weniger typisch sind (z. B. das Bewerten von Ergebnissen im Zuge der Einordnung in den Kontext oder das divergente Denken bei der Formulierung von Kontextfragen).

Die Abkürzungen der folgenden Aufgabenbeispiele A1 – A5 finden sich in Tabelle 12 im Anschluss an die Vorstellung der Aufgaben wieder. Tabelle 12 verdeutlicht die Einordnung der Aufgabenstellungen in das ausgearbeitete Modell.

Aufgabe	Aufgabentext	Einordnung
A1	<i>Um Fettflecke aus der Kleidung zu entfernen, verwendet man Reinigungsbenzin, während sich die Flecken mit Wasser nicht entfernen lassen. Erklären Sie zum einen die reinigende Wirkung von Reinigungsbenzin und zum anderen, warum sich Wasser nicht zur Entfernung von Fettflecken aus der Kleidung eignet!</i>	→ konvergentes Denken bei der Überprüfung der Ergebnisse vor dem Hintergrund eines Kontextes
A2	<i>Bei einer Vakuumdestillation sind nicht so hohe Temperaturen nötig wie bei einer Destillation unter Atmosphärendruck. Erklären Sie anhand des abgebildeten Phasendiagramms!</i>	→ Umgang mit grafischen Darstellungen bei der Durchführung von Untersuchungen

Das Diagramm zeigt ein Phasendiagramm mit der y-Achse 'Druck' und der x-Achse 'Temperatur'. Die Kurve ist in drei Bereiche unterteilt: 'Festkörper' (oben links), 'Flüssigkeit' (Mitte) und 'Gas' (unten rechts). Ein horizontaler Strich bei '1 atm' schneidet die Kurve an zwei Stellen: einer, die als 'Schmelzpunkt bei Normaldruck' beschriftet ist, und einer, die als 'Siedepunkt bei Normaldruck' beschriftet ist.

Eine genauere Betrachtung von Aufgaben aus Schulbüchern und Unterrichtsmaterialien zeigt, dass insbesondere der Bereich des divergenten Denkens vernachlässigt wird. Darüber hinaus soll dem Aspekt der Planung im Unterricht durch die Schülerinnen und Schüler stärkeres Gewicht zukommen. Das folgende Beispiel zeigt zwei Aufgabenformulierungen mit identischem Inhalt, die sich beide durch die Aufgabenanforderung „divergentes Denken“ auszeichnen, sich aber unterschiedlichen Kompetenzbereichen im Erarbeitungsprozess zuordnen lassen:

Aufgabe	Aufgabentext	Einordnung
A3	<i>Entwerfen Sie ein Demonstrationsversuch, mit dem die größere Flüchtigkeit von Winterbenzin im Vergleich zu Sommerbenzin auf einfache Weise gezeigt werden kann!</i>	→ divergentes Denken bei der Darstellung von Ergebnissen
A4	<i>Überlegen Sie sich ein Experiment, das die Annahme der unterschiedlichen Flüchtigkeiten von Sommer- und Winterbenzin belegt!</i>	→ divergentes Denken bei der Planung von Untersuchungen

Im ersten Beispiel müssen die Schülerinnen und Schüler anhand ihres Wissens über die größere Flüchtigkeit von Winterbenzin im Vergleich zum Sommerbenzin einen Versuch entwickeln, während den Schülerinnen und Schülern im zweiten das Ergebnis gewissermaßen nicht bekannt ist, so dass sie zunächst eine Hypothese formulieren müssen (z. B. Winterbenzin geht bei identischen Bedingungen leichter in die Gasphase über als Sommerbenzin), und um diese Annahme zu bestätigen oder zu widerlegen, müssen die Lernenden im nächsten Schritt ein geeignetes Experiment entwickeln.

Das folgende Aufgabenbeispiel wurde in leicht veränderter Form als Klausuraufgabe im Anschluss an eine Kontexteinheit „Treibstoffe“ eingesetzt. Diese Aufgabe zeichnet sich durch Komplexität aus, sie enthält verschiedene kognitive Anforderungen und lässt sich insgesamt dem Bereich der Bewertung und Einordnung der Ergebnisse in den Kontext zuordnen.

Aufgabe	Aufgabentext	Einordnung
A5	<p><i>Bei einem Gang über eine Tankstelle finden Sie deutlich sichtbare Warnhinweise, aber auch verborgene Schutz- und Vorsichtsmaßnahmen. Diese verschiedenen Maßnahmen können Sie mit den typischen Eigenschaften und der molekularen Struktur des Benzins, die Sie im vorangegangenen Unterricht kennen gelernt haben, begründen.</i></p> <p><i>Zu diesen Maßnahmen zählen:</i></p> <ul style="list-style-type: none"> • <i>„Rauchen verboten“-Schilder und deutlich sichtbar angebrachte Feuerlöscher</i> • <i>Tankkrüssel sind mit einer Saugvorrichtung versehen, die sicherstellt, dass alle Gase, die während des Tankvorgangs entweichen, aufgefangen und gesammelt werden.</i> • <i>Das auf Tankstellen ablaufenden Wasser (Regen, Wasser aus Waschanlagen) wird über Rohrleitungen in sog. Wasserrückgewinnungsanlagen aufgefangen und wiederaufbereitet. Außerdem sind Tankstellen mit einer absolut flüssigkeitsdichten Fahrbahn versehen, so dass keine Flüssigkeit in den Boden versickern kann.</i> • <i>Gefahrensymbol und Warnhinweise auf jeder Tanksäule</i> <p><i>Beschreiben Sie die genannten Maßnahmen mit eigenen Worten! Erklären Sie die Maßnahmen und gehen Sie dabei auf die molekulare Struktur und zwischenmolekularen Wechselwirkungen des Benzins ein!</i></p>	<p>→ Umgang mit Sprache, konvergentes Denken, Umgang mit Modellvorstellungen bei der Einordnung von Ergebnissen in einen Kontextzusammenhang</p>

Die folgende Tabelle 12 zeigt die Einordnung der genannten Aufgabenbeispiele in die einzelnen Felder des zusammengefügteten Analysemodells.

	divergentes Denken	konvergentes Denken	Umgang mit grafischen Darstellungen	Umgang mit Sprache	Umgang mit Zahlen	Bewerten	Umgang mit mentalen Modellen
Kontext erschließen, Fragen formulieren							
Fragen identifizieren, die chemisch untersucht werden können							
Vorgehensweise planen	A4						
Untersuchungen durchführen			A2				
Ergebnisse darstellen und kommunizieren	A3						
Ergebnisse bewerten und einordnen (Kontext)		A1, A5	A5	A5			A5
Ergebnisse bewerten und einordnen (Chemie)							
Abstraktion zu Basis-konzepten							

Tab. 12: Einordnung der Aufgabenbeispiele A1 – A5 in das Analysemodell.

4.3 Testaufgaben der Untersuchung

An diese theoretischen Ausführungen schließt sich die Beschreibung und Kategorisierung der Testaufgaben des eingesetzten Instruments zur Wissens- und Kompetenzerfassung der Hauptstudie an. Da sich die Klassifizierung der Testaufgaben sachlogisch an die bisherigen Ausführungen anschließt, werden sie bereits an dieser Stelle aufgegriffen. Die Zielstellung der Untersuchung wird ebenso wie die Fragestellungen im empirischen Teil der vorliegenden Arbeit in einem eigenen Kapitel betrachtet, an dieser Stelle soll es ausschließlich um die Betrachtung der Gestaltung und Anforderungen der eingesetzten Aufgaben gehen.

Ebenso werden die Testaufgaben der Pilotstudie an späterer Stelle vorgestellt, in diesem Zusammenhang werden auch die Veränderungen verdeutlicht, die einzelne Items des Pilottests für den Einsatz in der Hauptstudie erfahren haben.

In Kapitel 4.3 der vorliegenden Arbeit findet sich in verkürzter Form eine Zusammenstellung aller Aufgaben des Testinstruments der Hauptuntersuchung (vollständige Aufgabenstellungen siehe Anhang E, S. 187), die in dieser Tabelle eingeführten Kurzbezeichnungen der Aufgaben dienen der Erhöhung der Lesbarkeit des folgenden Textes. In einem nächsten Schritt wird anhand ausgewählter Aufgabenbeispiele aus dem Instrument eine Einordnung nach den beschriebenen Kategorien vorgenommen. Einen Gesamtüberblick über die Einordnung aller Testaufgaben des Hauptuntersuchungsinstruments geben im Anschluss daran die Tabellen 13a und 13b (Kap. 4.3.9, S. 68, 69).

4.3.1 Inhalt, Lehrplan und Richtlinien

Insgesamt sind die entwickelten Aufgaben in den Themenkomplex Säuren, Basen und Neutralisation eingebettet. Die Testaufgaben sind nach ihrer Oberflächenstruktur zu drei thematischen Blöcken geordnet. Den thematischen Rahmen für die Aufgaben des ersten Testteils liefert das Krankheitsbild Sodbrennen. Diesem Testteil sind fünf Teilaufgaben zugeordnet. Der mittlere Teil des Tests orientiert sich inhaltlich an der Frage nach der Wirkung und der Wirkungsweise von Zahnpflegetablets, wobei sich dieser Teil aus vier Aufgaben zusammensetzt. Die Aufgaben des letzten Testteils werden nicht durch eine übergeordnete, inhaltliche Themenstellung zusammengefasst, sondern sind vielmehr dadurch charakterisiert, dass sie sich unter dem Begriff „Schulwissen zu Säuren und Basen“ subsumieren lassen. Dieser Teil des Tests besteht aus sechs Aufgaben.

Während sich die ersten beiden Testteile durch einen Kontextrahmen beschreiben lassen (Sodbrennen und Zahnpflegetablet als Teilaspekte eines Kontextes wie z. B. Gesundheit oder Ernährung), ist der dritte Testteil als eine Abstraktion von dem Kontext zu betrachten. Bei diesen Aufgaben geht es um fachliche Ausprägungen des übergeordneten Basiskonzepts Donator-Akzeptor, die in der Säure-Base-Thematik von Bedeutung sind. Es handelt sich folglich um vom Kontext abstrahierte Fachinhalte bzw. um einen „innerfachlichen Kontext“ (vgl. dazu Steinhoff 2004). Der Kontext, in dem diese Fachinhalte erworben werden, ist der tatsächlich erlebte Chemieunterricht in der Schule.

Die Testaufgaben werden exemplarisch zur Erfassung einzelner Aspekte des Donator-Akzeptor-Konzepts entwickelt. Die Tiefenstruktur lässt sich zum einen aus fachlicher Sicht mit dem Säure-Base- und Neutralisationskonzept umschreiben, zum anderen kann die Tiefenstruktur durch die Aufgabenanforderungen näher erfasst werden, die für das naturwissenschaftliche Denken und Arbeiten charakteristisch sind.

Die Lehrplanvalidität der Aufgaben soll dahingehend berücksichtigt werden, dass von teilnehmenden Lehrkräften für jede einzelne Aufgabe eine Einschätzung vorgenommen wird, ob sie im Unterricht behandelt wurde und ob die Art der Aufgabenstellung generell für ihren Unterricht eine Bedeutung hat (dazu genauer Kap. 6.3.2).

Abschließend werden die Items des Testinstruments dahingehend klassifiziert, welche konkreten Wissensbestände für die erfolgreiche Bearbeitung der Aufgabe benötigt werden. Dabei werden folgende Wissensbestandteile unterschieden:

- qualitatives Verständnis zum pH-Wert, Kenntnisse zu Indikatoren (z. B. in Aufgabe „Kurve beschreiben“)

- Kenntnisse zur Stoffklasse Säuren und ihren Eigenschaften (z. B. in Aufgabe „Demonstrationsversuche“)
- Kenntnisse zur Stoffklasse Basen (z. B. in Aufgabe „basische Stoffe“)
- Kenntnisse zur Neutralisation (z. B. in Aufgabe „Reaktionsgleichung“)

Darüber hinaus gibt es eine Reihe von Testaufgaben, in denen keine der genannten Wissensbestände zur erfolgreichen Bearbeitung notwendig sind, wie z. B. in Aufgabe „Studie“.

4.3.2 Unterrichtsphasen

Eine konkrete Zuordnung zu Unterrichtsphasen vorzunehmen, ist nicht erforderlich, da das Instrument im Rahmen der Untersuchung in erster Linie zur Lernstandserhebung konstruiert wurde und daher im Bereich der Leistungsmessung eingeordnet werden muss. Dennoch lassen sich für einzelne Aufgaben Empfehlungen aussprechen, für welche Phase eines Unterrichts nach *Chemie im Kontext* sie geeignet erscheint (vgl. dazu Kap. 2.1, Abb. 1, S. 7 der vorliegenden Arbeit).

Die Aufgabe „Aspekte zur Medikamentenentwicklung“ eignet sich für den Einsatz in der Begegnungsphase, denn anhand dieser Fragestellung können die zahlreichen Perspektiven aufgezeigt werden, die bei der Auseinandersetzung mit dem Thema zum Tragen kommen. Die Aufgabe „Versuch zum Gas auffangen“ lässt sich der Phase der Planung zuordnen, wenn die Schülerinnen und Schüler überlegen, wie sich das Gas, das bei der Einnahme von Natriumhydrogencarbonat im Magen entsteht, auffangen lässt, um es genauer untersuchen zu können.

In die Phase der Durchführung lassen sich Aufgaben einordnen, in denen konkrete Ergebnisse zu Fragestellungen gewonnen werden, wie z. B. in den Aufgaben „Reaktionsgleichung“ und „Berechnung“. Für die letzte Phase der Vertiefung und Vernetzung eignet sich der überwiegende Teil der Testaufgaben. Zentral sind hier das Anwenden von Kenntnissen in bekannten und neuen Zusammenhängen und das Abstrahieren von Fachkenntnissen vom ursprünglichen Erarbeitungskontext, wie z. B. in den Aufgaben „Gegenmittel Sodbrennen“, „Fehlersuche“ und „Leitfähigkeit“.

4.3.3 Art der möglichen Lösungswege

Da es sich bei den Testaufgaben um ein Instrument handelt, das in einer verhältnismäßig großen Stichprobe eingesetzt wird, bieten die Aufgaben den Schülerinnen und Schülern kaum die Möglichkeit, die Art des Lösungsweges frei zu wählen, sondern die Wege sind vorgegeben. Es wird davon ausgegangen, dass sich das Problem der Mathematisierung in der Chemie nicht im gleichen Maße wie in dem Physikunterricht stellt, so dass eine Beschreibung der Lösungswege nach den vorgeschlagenen Kategorien (siehe Kap. 3.3.3.3 dieser Arbeit, Häußler et al. 1998a, Fischer

et al. 2002) auf die Chemie nicht unmittelbar übertragen werden kann. Für eine Unterscheidung der Lösungswege in der Chemie könnten experimentelle, stoffliche und modellhafte Betrachtungen geeignet sein, wobei diese drei Unterscheidungen nicht als qualitativ gleichwertig angesehen werden können (vgl. dazu Woest 2004).

Anhand von Aufgabe III.6 wird beispielhaft gezeigt, wie unterschiedliche Lösungswege aussehen könnten. In der Aufgabenformulierung wird kein konkretes Beispiel für die saure Lösung genannt. In einer möglichen Antwort zieht der Lernende eine bestimmte Säure heran, die er im Unterricht kennen gelernt hat und von der er weiß, dass eine wässrige Lösung dieser Säure den elektrischen Strom leitet (z. B. Salzsäure enthält H^+ -Ionen). Ein anderer Lösungsweg könnte so aussehen, dass der Begriff Säure als Begriff für ein Konzept, losgelöst von dem Stoff Säure verwendet wird (saure Lösungen enthalten H^+ -Ionen).

In Aufgabe III.4 sind die Lernenden aufgefordert, einen geeigneten Indikator für das Anzeigen des neutralen pH-Bereichs zu identifizieren. Indem die Angaben zu den Umschlagsbereichen verschiedener Indikatoren ausgespart werden, können verschiedene Lösungswege produziert werden: durch Versuche kann der richtige Indikator identifiziert werden, aber auch eine theoretische Vorgehensweise ist möglich (z. B. der neutrale pH-Wert liegt bei 7, der Umschlagsbereich von Bromthymolblau liegt laut Literatur bei $pH = 6,8 - 7,2$).

4.3.4 Antwortformate

In diesem Test kommen verschiedene Antwortformate zum Einsatz. In einer ersten übergeordneten Einteilung werden die Aufgaben nach der Offenheit ihrer Antworten unterschieden, wobei es sich bei einem offenen Format um Antworten handelt, die von den Schülerinnen und Schülern in Form von Text, Rechnung oder Zeichnung frei formuliert werden, und bei den geschlossenen Formaten um verschiedene Formen von Multiple-Choice-Aufgaben.

Zusammengefasst findet sich bei 13 der insgesamt 19 Items ein offenes, bei den übrigen sechs ein geschlossenes Antwortformat. Bei der Klassifizierung des offenen Antwortformats wird in fünf Kategorien unterschieden: zwei Aufgaben lassen sich nach ihrem erforderlichen Antwortformat den Stichwort-Antworten zuordnen (z. B. Aufgabe „Internet-Suchbegriffe“), sechs Aufgaben erfordern eine längere Textantwort (z. B. Aufgabe „Grafik beschreiben“), jeweils eine Aufgabe erfordern als Lösung eine Zeichnung (Aufgabe „Versuch zum Gas auffangen“), eine Rechnung (Aufgabe „Berechnung“) oder das Formulieren einer Reaktionsgleichung (Aufgabe „Reaktionsgleichung“). Bei den Multiple-Choice-Aufgaben werden einfache und komplexe Formate unterschieden, wobei fünf Aufgaben durch ein einfaches Format gekennzeichnet sind (z. B. Aufgabe „Studie“), d. h. eine richtige Antwort muss aus einer Liste von möglichen Antworten aus-

gewählt werden. In einem Fall kommt ein komplexes Multiple-Choice-Format zum Tragen, bei dem die Schülerinnen und Schüler für jede genannte Aussage entscheiden müssen, ob sie zutrifft oder nicht (Aufgabe „Kaugummiwirkung“).

4.3.5 Grad der Offenheit

Anhand dieser Kategorisierung wird festgestellt, ob und in welcher Intensität der Lösungsweg in einer Aufgabe festgelegt wird. In diesem Zusammenhang gilt wiederum, dass bei einem Instrument zur Lernstandserhebung bei einer relativ großen Stichprobe mit Blick auf die Auswertung keine große Offenheit gewährt werden kann. Nach der Kategorisierung nach Fischer et al. (2002) ist festzustellen, dass die Art des Lösungsweges in allen Fällen mit hoher Intensität (Stufe 3) vorgegeben ist, entweder direkt durch die Aufgabenformulierung wie z. B. bei Aufgabe „Versuch zum Gas auffangen“ durch „Erstelle eine Zeichnung... und beschrifte sie“ oder aber durch Hinzufügen von Linien, auf denen die Antwort formuliert werden soll, wie z. B. bei Aufgabe „Leitfähigkeit“. Bei dieser Aufgabe ist auch eine zeichnerische Lösung denkbar, aber die Existenz der Linien legt dem Probanden eine schriftliche Form der Erläuterung nahe.

4.3.6 Erforderliches Experimentierverhalten

Eine Betrachtung der Testitems in Hinblick auf das erforderliche Experimentierverhalten erübrigt sich weitgehend bei dem eingesetzten Testinstrument. Das Entwerfen eines Versuchs zum Auffangen eines Gases lässt sich anhand der vorgeschlagenen Kategorisierung nach Fischer et al. (2002) zwischen den beiden Stufen „organisierend“ und „konzeptuell“ einordnen, da Informationen zum theoretischen Hintergrund im Aufgabentext geliefert werden, d. h. die Schülerinnen und Schüler müssen zwar keine eigenen Hypothesen entwickeln, aber es werden keine Gerätschaften für den Versuchsaufbau vorgegeben. Durch entsprechende Änderungen in der Aufgabenstellung lassen sich die Anforderungen an das Experimentierverhalten der Lernenden variieren.

4.3.7 Anforderungsmerkmale

Die Anforderungsmerkmale, mit deren Hilfe die eingesetzten Testitems beschrieben werden, sind bereits vorgestellt worden (Kap. 3.3.3.7). Alle sieben Aufgabenanforderungen werden im Test thematisiert, im Folgenden wird jeder der sieben Anforderungen ein Testitem zugeordnet, die Zuordnung der restlichen Items kann den Tab. 13a und 13b (S. 68 und 69) entnommen werden. Die Anforderung des divergenten Denkens kommt z. B. in Aufgabe „Aspekte der Medikamentenentwicklung“ zum Tragen, die Anforderung des konvergenten Denkens z. B. in Aufgabe „basische Stoffe“, die des Umgangs mit grafischen Darstellungen z. B. in Aufgabe „pH-Kurven“

und mit Zahlen in der Aufgabe „Berechnung“. In Aufgabe „Fehlersuche“ ist die Anforderung des Bewertens charakterisierend, in der Aufgabe „Leitfähigkeit“ wird der Umgang mit Modellen thematisiert, die Fähigkeit des adressatengerechten Verbalisierens beschreibt die Anforderung in Aufgabe „Fragen formulieren“. Die Zuordnung eines einzelnen Anforderungsmerkmals zu einem Item erfolgt anhand einer Schwerpunktsetzung, wobei aber berücksichtigt werden muss, dass sich z. B. die Anforderung des Verbalisierens natürlich gerade durch den hohen Anteil von Antwortformaten, die eine Textantwort verlangen, in vielen Aufgaben als zweite bedeutende Aufgabenanforderung vorhanden ist.

4.3.8 Kompetenzbereiche, Kompetenzstufen

In diesem Zusammenhang erfolgt eine beispielhafte Zuordnung der Testitems zu den acht Erkenntnisschritten, die in Kapitel 2.2 und 4.1.1 formuliert und vorgestellt wurden. Die Kategorisierung aller Items anhand dieser Klassifizierung kann wiederum der Tabelle 13b (S. 69) entnommen werden.

Dem ersten Schritt der Kontexterschließung und des Fragenformulierens wird z. B. die Aufgabe „Aspekte der Medikamentenentwicklung zugeordnet“, dem zweiten der Identifizierung von zur Untersuchung geeigneter Fragen wird die Aufgabe „Fragen formulieren“ zugewiesen. Mit der Aufgabe „Internet-Suchbegriffe“ lässt sich der Bereich des Planens der Vorgehensweise beschreiben, mit Aufgabe „Berechnung“ das Durchführen von Untersuchungen. Der Ergebnisdarstellung wird die Aufgabe „Demonstrationsversuche“ zugeordnet, die Ergebnisbewertung innerhalb des Kontextrahmens thematisiert die Aufgabe „Kaugummiwirkung“. In Aufgabe „pH-Kurven“ geht es im Schwerpunkt um die Bewertung von Ergebnissen vor dem Hintergrund chemischer Fachkenntnisse. Eine Abstraktion zu Basiskonzepten, die den letzten Schritt im Erkenntnisprozess beschreiben, erfolgt in der Aufgabe „Leitfähigkeit“.

4.3.9 Überblick über die Kategorisierung aller Testaufgaben

Die beiden folgenden Tabellen 13a und 13b zeigen die Einordnung der Testaufgaben des Haupttests einerseits nach den eher formalen Kategorien Oberflächenstruktur, Wissensbestände und Antwortformat und verdeutlichen andererseits die Zuordnungen der Aufgaben zu den Kategorien Erkenntnisschritte und Aufgabenanforderungen. Eine Zuordnung nach Eignung für bestimmte Unterrichtsphasen, Art der möglichen Lösungswege, Grad der Offenheit der Lösungswege und erforderlichem Experimentierverhalten unterbleibt aus oben ausgeführten Gründen.

Die weiteren Merkmale und Kriterien, die bei der Gestaltung und Auswahl der Testitems des Untersuchungsinstruments in Hinblick auf die Hinweise zur Testkonstruktion zu berücksichtigen

sind, werden nicht in diesem Zusammenhang vertieft, sondern im Rahmen der Beschreibungen der Pilot- und Hauptstudie dargestellt (siehe Kap. 5.3.2 und 6.3.2).

Ausprägung Aufgabe	Oberflächenstruktur			Wissensbestände					Antwortformat						
	Sodbrennen	Kaugummi	Schulchemie	pH-Wert, Indikator	Säuren (Eig.)	Basen	Neutralisation	keine der angeg.	Begriffe, Stichwort	Text	Zeichnung	Rechnung	Reaktionsgl.	MC (einfach)	MC (kompl.)
„Gegenmittel Sodbrennen“	X						X							X	
„Aspekte der Medikamentenentwicklung“	X							X*	X						
„Fragen formulieren“	X							X*		X					
„Internet-Suchbegriffe“	X							X*	X						
„Versuch zum Gas auffangen“	X							X			X				
„Reaktionsgleichung“	X						X						X		
„Berechnung“	X							X				X			
„Informationen einholen“	X							X		X					
„Grafik beschreiben“		X		X						X					
„Grafik interpretieren“		X		X						X					
„Fehlersuche“		X		X						X					
„Kaugummivirkung“		X						X*							X
„Studie bewerten“		X						X*						X	
„basische Stoffe“			X			X								X	
„pH-Kurven“			X	X										X	
„Ratschlag“			X				X			X					
„Indikatoren“			X	X										X	
„Demonstrationsversuche“			X		X					X					
„Leitfähigkeit“			X		X					X					

Tab. 13a: Kategorisierung der Testaufgaben der Hauptuntersuchung anhand der vorgestellten Kriterien, Teil 1; für die mit einem Stern gekennzeichneten Aufgaben sind keine Kenntnisse notwendig, die dem Schulwissen Chemie zugeordnet werden. Sie lassen sich auf der Basis von „Alltagswissen“ bearbeiten.

Ausprägung Aufgabe	Erkenntnissschritte								Aufgabenanforderung						
	Kon- text ersch.	chem. Fragen	Pla- nung	Durch- führung	Ergeb- nisse	Ei- nordn. (K.)	Ei- nordn. (Ch.)	Basis- konzept	div. Denken	konv. Denken	Zahlen	Gra- phen	Spra- che	Model- le	Bewer- tung
„Gegenmittel Sodbrennen“						X				X					
„Aspekte der Medikament- entwicklung“	X								X						
„Fragen formulieren“		X											X		
„Internet- Suchbegriffe“			X							X					
„Versuch zum Gasauffangen“			X							X					
„Reaktionsglei- chung“					X								X		
„Berechnung“				X							X				
„Informationen einholen“			X							X					
„Grafik beschreiben“					X							X			
„Grafik interpretieren“						X				X					
„Fehlersuche“						X									X
„Kaugummi- wirkung“						X									X
„Studie bewerten“						X									X
„basische Stoffe“							X			X					
„pH-Kurven“								X				X			
„Ratschlag“								X		X					
„Indikatoren“								X		X					
„Demonstra- tionsversuche“					X					X					
„Leitfähigkeit“								X						X	

Tab. 13b: Kategorisierung der Testaufgaben der Hauptuntersuchung anhand der vorgestellten Kriterien, Teil 2.

Die Tabellen 13a und 13b spiegeln die Komplexität einer differenzierten Betrachtung von Aufgaben wider. An dieser Stelle wird zugleich deutlich, dass es kaum möglich ist, im Rahmen einer statistischen Analyse alle aufgeführten Facetten einer Aufgabenstellung zu untersuchen, sondern dass Schwerpunkte festgelegt werden müssen. Trotzdem werden im Rahmen der statistischen Untersuchung der Aufgaben ein Teil der aufgeführten Aspekte berücksichtigt, indem über die unterrichtende Lehrkraft Informationen z. B. zu den behandelten Unterrichtsthemen und zum Vermittlungsstand der in den Aufgaben erforderlichen Kenntnissen und Fertigkeiten eingeholt werden. Ein hoher Schwierigkeitsgrad einer Aufgabe kann z. B. darin begründet sein, dass den Lernenden der Inhalt unbekannt ist, dass die Aufgabenanforderungen zu hoch sind oder dass das Format unbekannt ist und daher Probleme bereitet.

5 Pilotstudie: Entwicklung und Erprobung von Testaufgaben zur Erfassung zentraler Kompetenzen im Chemieunterricht

Die durchgeführte Untersuchung im Rahmen dieser Arbeit setzt sich aus einer Pilot- und einer Hauptstudie zusammen. Ziel der Pilotstudie ist die Ausarbeitung und Erprobung von Testaufgaben, die in der Hauptstudie als zentrales Untersuchungsinstrument eingesetzt werden.

Im Folgenden werden die Ziele und Fragestellungen der Pilotuntersuchung vorgestellt. Im Anschluss daran wird die Stichprobe, mit der das Untersuchungsinstrument erprobt wurde, und die Aufgaben des Testinstruments beschrieben. Im Anschluss werden Ergebnisse zusammengestellt, und Schlussfolgerungen und Änderungen, die sich daraus für die Gestaltung des Testinstruments ergeben, aufgezeigt und diskutiert.

Tabelle 14 zeigt in verkürzter Form die einzelnen Arbeitsschritte für die Planung und Durchführung der Untersuchung.

<p>VORARBEITEN</p> <p>Schritt 1: Konstruktion eines Modells zu zentralen Kompetenzen in einem Unterricht nach <i>Chemie im Kontext</i></p> <p>Schritt 2: Erste Erprobung des Modells anhand einer Materialanalyse</p>
<p>PILOTSTUDIE</p> <p>Schritt 3: Konstruktion von Testaufgaben zum Thema Säure, Base und Neutralisation</p> <p>Schritt 4: Gesprächsvalidierung der Testaufgaben mit Experten</p> <p>Schritt 5: mündliche Erprobung der Testaufgaben, Gruppenbearbeitung der Aufgaben, Videoaufzeichnung</p> <p>Schritt 6: Fragebogeneinsatz zur statistischen Überprüfung der Testitems</p> <p>Schritt 7: Auswertung der Videoaufzeichnung, Ausdifferenzierung des Erwartungshorizonts</p> <p>Schritt 8: Auswertung des Fragebogens</p> <p>Schritt 9: Ergebnisdarstellung und Überarbeitung der Testitems, Differenzierung der Punktevergabe und des Erwartungshorizonts anhand der Ergebnisse</p>
<p>HAUPTSTUDIE</p> <p>Schritt 10: Einsatz des überarbeiteten Fragebogens</p> <p>Schritt 11: Auswertung und Ergebnisdarstellung</p>

Tab. 14: Kurzdarstellung der Planung und Durchführung der Untersuchung.

5.1 Ziele der Untersuchung

Ziel der Pilotstudie ist die Entwicklung und Erprobung von Testaufgaben für einen Wissens- und Kompetenztest, der das zentrale Untersuchungsinstrument der Hauptstudie darstellt. Diese Testaufgaben sollen dazu geeignet sein, den Lernstand von Schülerinnen und Schülern in Hinblick auf den für den Chemieunterricht wichtige Fähigkeiten zu erfassen. Die Kompetenzanforderungen der Aufgaben ergeben sich aus den theoretischen Ausführungen zu den Zielen und deren inhaltlichen Konkretisierungen in den vorangegangenen Kapiteln 2 und 4 zur Unterrichtskonzeption *Chemie im Kontext*.

In der Pilotstudie geht es in erster Linie um die Überprüfung der Praktikabilität des Testinstruments: Bearbeitungsdauer, Schwierigkeitsgrad und Trennschärfe der Aufgaben, Differenzierung des Auswertungsschemas und des Erwartungshorizonts. Darüber hinaus wird u. a. das Itemverständnis kommunikativ überprüft, indem die Testaufgaben von Schülerinnen und Schülern in Gruppen mündlich diskutiert und bearbeitet werden (Videoaufzeichnung). Mit Hilfe der statistischen Ergebnisse wird u. a. sichergestellt, dass im Test keine ungeeigneten Distraktoren verwendet werden oder dass bestimmte Aufgabenformulierungen nicht zu einer geschlechtsspezifischen Benachteiligung führen. Das Ergebnis dieser Pilotstudie sind überarbeitete Testaufgaben, die das zentrale Untersuchungsinstrument der Hauptstudie darstellen. Darüber hinaus werden erste Aussagen darüber getroffen, welche Aufgabenanforderungen und -inhalte für Schülerinnen und Schüler besondere Schwierigkeiten darstellen.

5.2 Fragestellungen

Aus den dargestellten Zielen der Pilotstudie lassen sich folgende, übergeordnete Fragestellungen ableiten:

Fragestellung 1:

Sind die Aufgaben für die bearbeitenden Schülerinnen und Schüler in ihren Formulierungen und Anforderungen verständlich?

Fragestellung 2:

Welche Erkenntnisse liefert eine statistische Itemanalyse?

Im Folgenden werden die einzelnen Fragestellungen konkretisiert. Zur Bearbeitung der übergeordneten Fragestellungen werden in den folgenden Abschnitten untergeordnete Fragestellungen entwickelt und erläutert. Die Pilotphase untergliedert sich in zwei Phasen:

Phase 1: mündliche und schriftliche Bearbeitung der Testaufgaben,

Videoaufzeichnung der Bearbeitung in Gruppen

Phase 2: Fragebogeneinsatz

Die Fragestellungen gelten für beide Phasen der Fragebogentestung gleichermaßen, die Gewichtung der Beantwortung der Fragestellung unterscheidet sich jedoch, z. B. können aufgrund von Aussagen der bearbeitenden Schülerinnen und Schüler aus Phase 1 zum Schwierigkeitsgrad der Aufgaben nur Hinweise gewonnen werden, Aufschluss über den tatsächlichen Schwierigkeitsgrad der einzelnen Aufgabe vermag erst die statistische Überprüfung zu geben (Phase 2).

Fragestellung 1: Aufgabenverständnis der Schülerinnen und Schüler

Mit dieser Fragestellung wird geprüft, was die bearbeitenden Schülerinnen und Schüler unter den Inhalten und Anforderungen der einzelnen Testaufgaben verstehen. Auf diese Weise werden unklare und unverständliche Formulierungen in den Aufgabentexten identifiziert. Bei Aufgaben mit geringer Lösungshäufigkeit wird darauf geachtet, ob sich durch die inhaltliche Analyse der mündlichen Bearbeitung und der schriftlichen Antworten der Aufgaben Hinweise auf die Gründe für Schwierigkeiten oder die unterschiedliche Auffassung der Anforderungen finden lassen.

Fragestellung 1.1:

Sind die Aufgabeninhalte für Lernende verständlich formuliert?

Fragestellung 1.2:

Lassen sich Aufgabenformate ausmachen, die den Lernenden besondere Probleme bereiten?

Fragestellung 2: Erkenntnisse aus der Itemanalyse

Aus der Auswertung des Fragebogens lassen sich wichtige Informationen über die Qualität der einzelnen Items gewinnen. Anhand verschiedener Kriterien sollen problematische Testaufgaben ausfindig gemacht werden, zu diesem Zweck werden verschiedene statistische Gütekriterien herangezogen.

Fragestellung 2.1:

Reliabilität des Gesamttests: Welche Trennschärfen besitzen die einzelnen Items?

Fragestellung 2.2:

Lassen sich bei den Multiple-Choice-Aufgaben problematische Distraktoren finden, die die Probanden z. B. unbeabsichtigt „in die Irre führen“?

Fragestellung 2.3:

Gibt es unter den Testitems einzelne Aufgabe, die in besonderer Weise Probanden geschlechtsspezifisch benachteiligen?

Fragestellung 2.4:

Wie sieht die Schwierigkeitsverteilung der Items im Gesamtttest aus?

Fragestellung 2.5:

Sind die Testaufgaben geeignet, die Leistungen der Schülerinnen und Schüler hinreichend differenziert darzustellen?

5.3 Methode

Im Folgenden wird zunächst auf die Stichproben der Videoaufzeichnung und des Fragebogeneinsatzes eingegangen. Daran schließt sich eine inhaltliche Beschreibung des in der Pilotstudie eingesetzten Testinstruments an. Abschließend wird beispielhaft das Bewertungsverfahren der Testaufgaben und in diesem Zusammenhang auch die Ausformulierung des Erwartungshorizonts für die Testaufgaben vorgestellt.

Im Anschluss an die Expertenvalidierung der entwickelten Aufgaben wurde ein erster Testlauf des Fragebogens mit sechs Schülerinnen und Schülern einer Hamburger Gesamtschule durchgeführt: Er diente dazu sicherzustellen, dass bei den Testaufgaben keine schwerwiegenden, vermeidbaren Verständnisprobleme auftreten und dass die Bearbeitung insgesamt nicht mehr als eine Schulstunde in Anspruch nimmt. Die eigentliche Pilotstudie gliedert sich in zwei Teile: Zum einen wurden die Testaufgaben in vier Gruppen mit jeweils vier Schülerinnen und Schülern einer zehnten Klasse eines Kieler Gymnasiums mündlich erprobt, wobei diese Bearbeitung auf Video aufgezeichnet wurde. Zum anderen wurde der Fragebogen von 120 Schülerinnen und Schülern aus Schleswig-Holstein und Niedersachsen bearbeitet. Die Videoaufzeichnung und das Versenden der Fragebögen fand im gleichen Zeitraum statt, so dass die Erkenntnisse der Gruppenbearbeitung für die Aufgabengestaltung auch erst für den Fragebogen der Hauptstudie zum Tragen kommen und nicht bereits für den Fragebogeneinsatz der Pilotstudie berücksichtigt werden konnten. Hinweise aus der Gruppenbearbeitung der Testaufgaben für die Überarbeitung der Musterlösungen des Erwartungshorizonts wurden jedoch bereits in der Pilotstudie berücksichtigt. Für die Bearbeitung der Fragebögen in Kleingruppen, die auf Video aufgezeichnet wurden, wurden in der Klasse bei einer ersten Besprechung des Vorhabens drei Tage vor der eigentlichen Aufzeichnung bereits Vierergruppen gebildet. Da nur zwei Kameras zur Verfügung standen, bearbeiteten die Gruppen in zwei aufeinander folgenden Zeitabschnitten die Aufgaben. Die Aufzeichnung einer Gruppenarbeit kann in der Auswertung kaum berücksichtigt werden, da am

Schulgebäude bauliche Maßnahmen durchgeführt wurden, die in der Lautstärke die Schülergespräche überdeckten. In einer Einführung wurden die Gruppen gebeten, die Aufgaben des Fragebogens zunächst mündlich zu bearbeiten und zu diskutieren und dann eine Konsenslösung der Gruppe auf dem Fragebogen zu notieren. Für die Aufgabenbearbeitung hatten die Gruppen 30 Minuten Zeit zur Verfügung.

Die Pilotstudie fand im April/Mai 2004 statt, im Juni wurden die Daten ausgewertet und die Testaufgaben und Musterlösungen anhand der Ergebnisse überarbeitet, um dann zum Ende des Schuljahres 2003/04 bzw. zu Beginn des Schuljahres 2004/05 in verschiedenen Klassen eingesetzt zu werden.

Der Fragebogeneinsatz der Hauptstudie erstreckt sich über einen Zeitraum von mehreren Monaten, weil je nach Lage der Sommerferien der Fragebogen noch vor oder erst nach den Ferien eingesetzt wurde.

5.3.1 Stichprobenbeschreibung

Stichprobe 1 (Videoaufzeichnung)

An der Videoaufzeichnung hat eine zehnte Klasse eines Kieler Gymnasiums zum Ende des Schuljahres 2003/04 teilgenommen. Diese Klasse steht in keinerlei Verbindung zum Projekt *Chemie im Kontext*, die unterrichtende Lehrkraft hat ebenfalls keinerlei Erfahrung mit der Umsetzung der Konzeption. Da der Schwerpunkt in der Pilotphase auf der Testung des eigentlichen Instruments liegt, ist die konzeptuelle Orientierung des erlebten Chemieunterrichts unerheblich. Ausschlaggebend ist vielmehr, dass sich die Klasse bereits mit der Säure-Base-Thematik auseinandergesetzt hat. An der Videoaufzeichnung nehmen vier Gruppen mit je vier Schülerinnen und Schülern teil, zur Bearbeitung der Aufgaben standen den Schülergruppen etwa 30 Minuten zur Verfügung.

Stichprobe 2 (Fragebogenbearbeitung)

An der Befragung mittels Fragebogen nahmen insgesamt 120 Schülerinnen und Schüler aus sechs Klassen teil. Die teilnehmenden Klassen stammen mit einer Ausnahme aus Schleswig-Holstein, eine teilnehmende Klasse stammt aus Niedersachsen. Es handelt sich bei den Lerngruppen um vier 10. Klassen aus Gymnasien, um eine aus 10. und 11. Klasse zusammengesetzten Lerngruppen aus einer Gesamtschule und einer 10. Klasse aus einer Realschule. Die Geschlechterverteilung insgesamt sieht in den sechs teilnehmenden Klassen folgendermaßen aus: 73 Schüler und 47 Schülerinnen nehmen an der Studie teil. Das Durchschnittsalter der Probanden liegt bei 16,1 Jahren. Etwa ein Drittel der Probanden (37,5 %) ist entschlossen, Chemie nach

der 10. Klasse abzugeben, in der gleichen Größenordnung liegt der Anteil der Schülerinnen und Schüler, die sich noch nicht entschieden haben (33 %), nur ein Prozentsatz von 5,7 % möchte Chemie als Leistungskurs in der Oberstufe besuchen, etwa ein Viertel (23,9 %) entscheidet sich für Chemie als Grundkurs in der Oberstufe. In allen Lerngruppen ist die Säure-Base-Thematik bereits behandelt worden, der Zeitpunkt der Erhebung lag in allen Fällen am Ende des Schuljahres 2003/04 im Mai/Juni 2004. In fünf der insgesamt sechs teilnehmenden Klassen wurde die Säure-Base-Thematik in einem Kontextunterricht bearbeitet (die zusammengesetzte Gesamtschulklasse stellt eine Ausnahme dar), alle Klassen hatten bis zu dem Zeitpunkt der Einheit keinerlei Erfahrung mit der Konzeption *Chemie im Kontext*. Folglich handelt es sich für alle fünf Klassen um die erste Unterrichtseinheit nach *Chemie im Kontext*. Bei den unterrichtenden Lehrkräften handelt es sich mit einer Ausnahme um Setlehrkräfte aus dem ersten bzw. zweiten Schulset aus Schleswig-Holstein und aus dem zweiten niedersächsischen Schulset (vgl. Kap. 2.3). Bei der sechsten Lehrkraft handelt es sich um ein Mitglied der *Chemie-im-Kontext*-Projektgruppe am IPN. Auf Seiten der Lehrkräfte ist also mehr oder weniger Erfahrung mit der Konzeption *Chemie im Kontext* vorhanden. Die Auswahl der teilnehmenden Klassen fand entsprechend nicht zufällig statt, sondern es wurden gezielt Projektlehrkräfte angesprochen.

5.3.2 Erhebungsinstrument der Pilotstudie

Im folgenden Abschnitt wird das Untersuchungsinstrument, das in dieser Pilotstudie auf seine Eignung hin untersucht wird, in seinem Aufbau, seinen Aufgabeninhalten und Aufgabenformaten genau beschrieben. Darauf folgt eine Einordnung der einzelnen Testitems anhand der Kriterien, die in Kapitel 4 erarbeitet und vorgestellt wurden.

Der Test richtet sich an Lernende am Ende der zehnten Klasse. Zu diesem Zeitpunkt entscheiden sie in der Regel, ob sie ihre Kenntnisse im Fach Chemie in der Oberstufe weiter vertiefen werden. Als fachliches Thema wurde die Säure-Base-Thematik gewählt, weil dieses Thema in der zehnten Klasse Unterrichtsgegenstand ist (z. B. schleswig-holsteinischer Lehrplan für die Sekundarstufe 1 Thema 3 (Klassenstufe 10) „Säuren und Basen“). Zudem sind zu diesem Thema in den Schulsets z. T. Kontexteinheiten erarbeitet und erprobt wurden, darüber hinaus bietet die Thematik viele Anknüpfungspunkte an die Lebenswelt der Lernenden.

Der Test der Pilotstudie enthält 16 Aufgaben, von denen es sich bei der ersten um eine „Aufwärmübung“ oder „Eisbrecher-Aufgabe“ (Häußler et al. 1998b) handelt, die im Folgenden nicht weiter in die Testauswertung eingebunden ist und in der statistischen Analyse keine weitere Beachtung findet. Zudem werden zwei Aufgaben in je zwei Teilaufgaben zerlegt, so dass insge-

samt 17 Items in die Auswertung einfließen. Der gesamte Test kann auf drei Wegen strukturiert werden:

- (1) Eine erste Orientierung liefern Aspekte zur konkreten Gestaltung der Aufgabe, dabei setzt sich diese erste Strukturierung aus Oberflächenstruktur und Antwortformat zusammen.
- (2) Die zweite Strukturierung wird inhaltlich durch die zur korrekten Lösung notwendigen Wissensbestände vorgegeben.
- (3) Die letzte und zugleich zentrale Strukturierung des Tests ergibt sich aus der Zuordnung der Aufgaben zu den in den Kapiteln 4.1.1 und 4.1.2 formulierten Aufgabenanforderungen und Erkenntnisschritten.

Eine Zuordnung aller Testaufgaben anhand der im Folgenden genauer ausgeführten Kategorien kann den Tabellen 39a und 39b im Anhang C der Arbeit (S. 179, 180) entnommen werden, sie entspricht weitgehend der Zuordnung der Testaufgaben des Fragebogens der Hauptuntersuchung (vgl. Kap. 4.3.9).

Die Testaufgaben sind vor ihrem Einsatz in einer ersten Überarbeitung im Gespräch mit verschiedenen Experten (Fachdidaktikern und Lehrkräften) in Hinblick auf ihren Inhalt und ihre Gestaltung validiert worden. Die Einordnung der Testaufgaben, die in der Hauptuntersuchung eingesetzt werden, anhand der vorgestellten Kategorien zur Klassifizierung von Aufgaben ist bereits in Kapitel 4.3 ausgeführt worden. Für die Aufgaben der Pilotphase wird diese Zuordnung daher in verkürzter Form dargestellt. In der Darstellung wird jeweils auf das Kapitel im theoretischen Teil der vorliegenden Arbeit verwiesen, welches die genauen Ausführungen zur Klassifizierung der Testitems aus der Hauptuntersuchung enthält. Eine Kurzübersicht über die eingesetzten Testaufgaben zeigt Tabelle 15, eine vollständige Version des Fragebogens findet sich im Anhang A (S. 169).

Aufgabe	Kurzbezeichnung	Kurzbeschreibung
I.1	„Gegenmittel Sodbrennen“	Aus einer Liste von fünf Substanzen müssen die Probanden die Substanz auswählen, die ihnen zur Bekämpfung von Magenübersäuerung geeignet erscheint.
I.2	„Aspekte der Medikamentenentwicklung“	Die Probanden sind aufgefordert, fünf Aspekte zu nennen, die ihnen bei der Entwicklung eines neuen Medikamentes wichtig erscheinen.
I.3	„Fragen formulieren“ „Internet-Suchbegriffe“	Die Probanden sind aufgefordert, je eine Frage zu formulieren, die sie ihrer Chemie- und Biologielehrkraft und einem Apotheker/einer Apothekerin stellen würden, um sich genauer über ein Medikament gegen Sodbrennen zu informieren. Die Probanden sind aufgefordert, drei Suchbegriffe zu formulieren, mit denen sie eine Internetrecherche zu einem Medikament gegen Sodbrennen beginnen würden.
I.4	„Versuch zum Gas auffangen“	Die Probanden sind aufgefordert, eine Versuchszeichnung zu einem Demonstrationsversuch zu erstellen, mit dem sie das entstehende Gas bei der Reaktion von NaHCO_3 mit saurer Lösung auffangen könnten.
I.5	„Reaktionsgleichung“ „Berechnung“	Die Probanden sind aufgefordert, die Reaktionsgleichung für die Neutralisationsreaktion von NaHCO_3 mit HCl zu formulieren. Die Probanden sind aufgefordert zu berechnen, welche Menge an HCl (in g) mit einer Tablette neutralisiert werden können.
II.1	„Grafik beschreiben“	Die Probanden sind aufgefordert, eine graphische Darstellung zur pH-Wert-Veränderung im Mund mit und ohne Kauen eines Zahnpflegekaugummis mit eigenen Worten zu beschreiben.
II.2	„Fehlersuche“	Die Probanden sind aufgefordert, die graphische Darstellung auf ihre Richtigkeit hin zu überprüfen und Fehler in der Darstellung zu identifizieren.
II.3	„Kaugummiwirkung“	Die Probanden sind aufgefordert, für vier Aussagen zur Wirkungsweise von Zahnpflegekaugummis jeweils zu entscheiden, ob die Aussage ein Argument für das Kaugummikauen ist.
II.4	„Studie bewerten“	Die Probanden sind aufgefordert, die Aussage einer beschriebenen Studie zur Wirksamkeit von Zahnpflegekaugummis zu bewerten, indem sie sich für eine von vier dargestellten Aussagen der Studie entscheiden.
III.1	„basische Stoffe“	Aus vier verschiedenen Aufzählungen von je drei Stoffen müssen die Probanden die Aufzählung auswählen, in der ausschließlich Stoffe genannt werden, die in Wasser gelöst basisch reagieren.
III.2	„pH-Kurven“	Aus fünf verschiedenen pH-Wert-Auftragungen müssen die Probanden die Darstellung auswählen, die die pH-Wert-Änderung bei einer Titration einer Salzsäurelösung mit einer gleich konzentrierten Natronlauge korrekt wiedergibt.
III.3	„Ratschlag“	Die Probanden sind aufgefordert, sich einem Ratschlag anzuschließen und diese Entscheidung zu begründen, wie mit verschütteter Säurelösung umgegangen werden könnte.
III.4	„Indikatoren“	Aus einer Liste von drei Indikatoren und deren pH-Umschlagsbereichen müssen die Probanden den Indikator auswählen, mit dem die pH-Neutralität von Regenwasser überprüft werden kann.
III.5	„Demonstrationsversuche“	Die Probanden sind aufgefordert, in tabellarischer Form zwei Eigenschaften von sauren Lösungen zu nennen und für jede der beiden Eigenschaften ein Versuch anzugeben, mit dem die genannte Eigenschaft demonstriert werden kann.
III.6	„Leitfähigkeit“	Die Probanden sind aufgefordert, mit Hilfe der Teilchenebene zu erklären, warum saure Lösungen den elektrischen Strom gut leiten, destilliertes Wasser hingegen den elektrischen Strom praktisch nicht leitet.

Tab. 15: Kurzbeschreibung der eingesetzten Testaufgaben der Pilotstudie.

Erste Strukturierungsebene

Während sich die ersten beiden Testteile durch einen Kontextrahmen beschreiben lassen (Sodbrennen und Zahnpflegekaugummi als Teilaspekte eines Kontextes wie z. B. Gesundheit oder Ernährung), ist der dritte Testteil als eine Abstraktion vom lebensweltlichen Kontext zu betrachten. Bei den Aufgaben des dritten Testteils geht es um fachliche Ausprägungen des übergeordneten Basiskonzepts Donator-Akzeptor, die in der Säure-Base-Thematik von Bedeutung sind, der inhaltliche Rahmen des letzten Testteils wird im Folgenden als „innerfachlicher Kontext“ bezeichnet (vgl. Kap. 4.3.1).

In diesem Test werden zudem verschiedene Aufgabenformate realisiert, diese Formate sollen durch das Antwortformat, das eine einzelne (Teil-) Aufgabe erfordert, klassifiziert werden. In einer ersten groben Einteilung können die Aufgaben nach der Offenheit ihrer Antworten unterschieden werden, dabei findet sich bei elf Items ein offenes Antwortformat, bei den übrigen sechs ein geschlossenes. Bei den geschlossenen Aufgaben handelt es sich um Multiple-Choice-

Aufgaben, in fünf Fällen mit einfachem, in einem Fall mit komplexem Format. Bei den offenen Aufgaben sind die Antwortformate vielfältiger: Von Aufzählungen von Begriffen (zwei Aufgaben) bis hin zu ausformulierten (Teil-) Sätzen (sechs Aufgaben) werden auch Antworten in Form einer Rechnungen, einer Reaktionsgleichung oder einer Versuchszeichnung (jeweils eine Aufgabe) erfordert (vgl. Kap. 4.3.4).

Zweite Strukturierungsebene

Obwohl es bei diesem Test nicht so sehr um eine Wissensabfrage, sondern stärker um das Erfassen der beschriebenen Kompetenzen mit ihrer Inhalts- und Handlungsdimension geht, werden die erforderlichen Wissensbestände der Aufgaben dargestellt (vgl. Kap. 4.3.1): Für drei Aufgaben benötigen die Probanden ein rein qualitatives Verständnis des pH-Wertes, zudem sollte der Begriff und der Zweck des Einsatzes eines Indikators bekannt sein. Ein weiteres Wissensgebiet, das den Schülerinnen und Schülern bekannt sein muss, um einzelne Aufgaben erfolgreich bearbeiten zu können, ist das der Stoffklasse der Säuren und einige ihrer typischen Eigenschaften (zwei Aufgaben). Für eine Aufgabe ist die Kenntnis der Stoffklasse der Basen zentral. Ein Verständnis für Neutralisationsreaktionen benötigen die Schülerinnen und Schüler beim Bearbeiten von drei Aufgaben. Nicht überraschend ist, dass die Probanden für die Bearbeitung jeder der Aufgaben aus dem letzten Testteil „Schulchemie“ einen der genannten Wissensbestände benötigen, während sich ein Großteil der restlichen Aufgaben aus den ersten beiden Testteilen ohne einen der genannten Wissensbestände aus der Chemie erfolgreich bearbeiten lässt.

Dritte Strukturierungsebene

Der Schwerpunkt des Tests und damit auch in der Beschreibung der Aufgaben liegt auf dem Erfassen der in Kapitel 4.1 vorgestellten, aus den Zielen der Konzeption *Chemie im Kontext* hergeleiteten zentralen Anforderungen, auf diese Weise wird somit die Tiefenstruktur der Aufgaben erfasst (vgl. Kap. 4.3.7).

5.3.3 Auswertung der Testitems

Im folgenden Abschnitt wird das Vorgehen bei der Bewertung und der Punktevergabe bei den Testaufgaben vorgestellt. Dazu wird zunächst für jede Aufgabe eine Musterlösung festgelegt, die alle Komponenten für eine vollständige und korrekte Lösung enthält (vollständiger Erwartungshorizont, siehe Anhang B, S. 178). Anhand von möglichen und erwarteten Schülerantworten wird dann in einem zweiten Schritt die Punktevergabe entschieden. Wichtige Aufschlüsse

zum Erwartungshorizont in Hinblick auf mögliche Antworten der Schülerinnen und Schüler liefern die Aussagen in der Gruppenbearbeitung (Videoaufzeichnung, Phase 1).

Punktevergabe

Anhand der erwarteten Antworten der Probanden wird für jede einzelne Aufgabe entschieden, ob sie dichotom oder mehrstufig bewertet werden soll, dabei gibt es maximal eine Abstufung für teilrichtige Lösungen. Insgesamt werden die Testaufgaben nach (1) vollständig und korrekt, (2) teilrichtig und (3) überwiegend falsch oder falsch bewertet. Insbesondere für die offenen Aufgaben, für die teilrichtige Antworten zugelassen sind, aber auch für die übrigen Aufgaben mit offenem Antwortformat ist die Formulierung eines detaillierten Erwartungshorizonts für die einzelnen Bewertungsstufen für die anschließende Bewertung zentral (Duit et al. 2001, Kap. 3.5 der vorliegenden Arbeit).

In dem Test der Pilotstudie werden acht der 17 Items dichotom mit richtig/falsch bewertet, bei den restlichen neun Items werden teilrichtige Lösungen zugelassen und gehen somit als mehrstufige Antworten in die Auswertung ein. Bei den Aufgaben mit mehrstufigem Antwortformat (die Richtigkeit der Aufgaben ist abgestuft in überwiegend richtig und vollständig richtig) handelt es sich in einem Fall um eine Multiple-Choice-Aufgaben, darüber hinaus werden neben sechs Aufgaben, die eine Stichpunkt- oder Textantworten erfordern, die Aufgaben „Versuch zum Gasanfassen“ und „Berechnung“ mit teilrichtigen Bearbeitungen gewertet. Die Verteilung der dichotomen und Mehrfachantworten auf die drei Testteile wird in der folgenden Tabelle 16 veranschaulicht.

Während im ersten Testteil „Sodbrennen“ überwiegend ein mehrstufiges Antwortformat vorherrscht, gibt es im zweiten Teil des Tests „Zahnpflegekaugummi“ ausschließlich dichotome Items. Im letzten Teil „Schulchemie“ liegt eine ausgewogene Verteilung der Antwortstufungen vor. Diese unterschiedliche Verteilung wird in der Auswertung durch eine entsprechende Gewichtung der einzelnen Aufgaben berücksichtigt.

Testteil	Aufgabe	dichotomes Antwortformat	mehrstufiges Antwortformat
Sodbrennen	„Gegenmittel Sodbrennen“		x
	„Aspekte der Medikamentenentwicklung“		x
	„Fragen formulieren“		x
	„Internet-Suchbegriffe“		x
	„Versuch zum Gas auffangen“		x
	„Reaktionsgleichung“	x	
	„Berechnung“		x
Kaugummi	„Grafik beschreiben“	x	
	„Fehlersuche“	x	
	„Kaugummiwirkung“	x	
	„Studie bewerten“	x	
Schulchemie	„basische Stoffe“	x	
	„pH-Kurven“	x	
	„Ratschlag“		x
	„Indikatoren“	x	
	„Demonstrationsversuche“		x
	„Leitfähigkeit“		x

Tab. 16: Verteilung der Antwortformate der eingesetzten Testaufgaben.

Erwartungshorizont

Für jede Aufgabe wird ausführlich festgelegt, welche Inhalte für eine vollständige und korrekte Antwort enthalten sein müssen. Bei neun Items werden teilrichtige Lösung in die Auswertung miteinbezogen, bei einem dieser neun handelt es sich um eine Multiple-Choice-Aufgabe, für die übrigen acht wird ebenfalls genau festgelegt, welche Komponenten eine Antwort enthalten muss, um als vollständig korrekt bzw. teilrichtig eingestuft zu werden. Die vollständige Ausformulierung des Erwartungshorizonts der Testitems befindet sich im Anhang dieser Arbeit (Anhang B, S. 178). An dieser Stelle wird am Beispiel der Aufgabe „Versuch zum Gasauffangen“ das Vorgehen veranschaulicht.

Eine Versuchsapparatur, die dazu geeignet ist, das entstehende Kohlenstoffdioxid aufzufangen, um damit weitere Versuche durchführen zu können, kann sehr unterschiedlich aussehen und auch unterschiedliche Kompliziertheitsgrade beinhalten. Die Komplexität der einzelnen Schülerzeichnungen wird bei der Auswertung nicht weiter berücksichtigt, sondern für das Erhalten der vollen Punktzahl müssen folgende Kriterien erfüllt sein:

- Die Zeichnung gibt einen sinnvollen Versuchsaufbau wieder, der in der Praxis funktionieren würde.
- Aus der Zeichnung geht hervor, worin sich das Reaktionsgemisch befindet (Zeichnung eines Behältnisses), wie das entstehende Gas abgeleitet (z. B. Glasrohr, Schlauch) und aufgefangen wird (z. B. in einem Kolbenprober, Luftballon, Reagenzglas).
- Wichtig ist, dass das Gas so aufgefangen wird, dass damit weitere Versuche durchgeführt werden können, d. h. das Gas darf z. B. nicht direkt in Bariumhydroxidlösung eingeleitet werden. Es ist jedoch nicht erforderlich, dass das entstehende Kohlenstoffdioxid in reiner

Form aufgefangen wird, für die weiteren Versuche (z. B. Nachweisreaktionen) wäre ein Gemisch mit Luft ausreichend.

- Die eingesetzten Chemikalien und zentralen Geräte im Versuchsaufbau sind ausreichend beschriftet.
- Die Videoaufzeichnung zeigt, dass bei den Schülerinnen und Schülern Unsicherheit darüber besteht, ob Kohlenstoffdioxid leichter oder schwerer als Luft ist. Da das Vorhandensein des Wissens über die unterschiedlichen Dichten von Luft und Kohlenstoffdioxid für die Auswertung der Aufgabe nicht von Interesse ist, wird in jedem Fall zu Gunsten des Aufgabenlösers entschieden: im ersten Fall wird unterstellt, dass er weiß, dass Kohlenstoffdioxid schwerer als Luft ist (seine Apparatur funktioniert in der Praxis), im zweiten Fall kann er nicht gewusst haben, dass Kohlenstoffdioxid schwerer als Luft ist, seine Apparatur würde in der Praxis folglich nicht funktionieren. In beiden Fällen wird die Lösung aber als richtig angesehen.

Für das Erhalten der Punktzahl für eine teilrichtige Lösung werden folgende Kriterien festgelegt:

- Die oben genannten Kriterien für eine vollständige Lösung müssen im Großen und Ganzen erfüllt sein, aber es können einzelne Schwächen auftreten. Als teilrichtig gelöst gilt die Aufgabe, wenn bei korrekter Zeichnung das entstehende Gas direkt in einer Nachweisreaktion weiterreagiert (z. B. Einleiten in Bariumhydroxidlösung, Einleiten in ein Gefäß mit brennender Kerze).
- Wenn aus der Zeichnung der Gedanke des Auffangens richtig hervorgeht, dann gilt die Aufgabe auch dann als teilrichtig gelöst, wenn ungenau gezeichnet wurde oder die Zeichnung nicht beschriftet ist.

Für den Fall, dass eine geschlossene Apparatur gezeichnet wurde, gab es keine Punkte, auch wenn alle anderen Kriterien erfüllt waren.

Durch das Einbeziehen teilrichtiger Lösungen in die Auswertung sollen vor allem typische gedankliche Fehler beim Lösen einer Aufgabe sichtbar gemacht werden. In Aufgabe „Ratschlag“ wird der Aufgabenlöser aufgefordert, sich begründet einem Ratschlag anzuschließen, wie mit der im Abzug verschütteten Säure verfahren werden soll. Durch die Wertung von „Zugabe von Natronlauge, weil sie die Säure neutralisiert“ als teilrichtige Antwort ist es möglich, den Prozentsatz der Schülerinnen und Schüler festzustellen, die ihr theoretisches Vorgehen nicht bis zum Ende durchdenken und daher nicht auf das Problem stoßen, dass sie die Menge der verschütteten Säure nicht kennen und daher auch bei bekannter Konzentration nicht bestimmen könnten, wie viel Lauge sie genau zugeben müssten, um eine neutrale Lösung zu erhalten.

5.4 Ergebnisse

In der folgenden Ergebnisdarstellung werden zunächst die Erkenntnisse zu allen genannten Fragestellungen aus der aufgezeichneten Gruppenbearbeitung dargestellt (Phase 1), im Anschluss daran wird auf die statistische Auswertung der Fragebogendaten (Phase 2) eingegangen. Die dargestellten Erkenntnisse werden z. T. durch mündliche und schriftliche Äußerungen der Schülerinnen und Schüler veranschaulicht.

5.4.1 Phase 1 der Pilotuntersuchung: Videoaufzeichnung der Gruppenbearbeitung

Fragestellung 1: Aufgabenverständnis der Schülerinnen und Schüler

Fragestellung 1.1

Sind die Aufgabeninhalte für Lernende verständlich formuliert?

Aus der Videoaufzeichnung geht aus einzelnen Äußerungen der Schülerinnen und Schüler bei der Aufgabenbearbeitung hervor, ob sie den Aufgabentext verstehen oder nicht. In der Aufgabe I.4. stellt möglicherweise der Begriff „Modellversuch“ eine Unklarheit dar, denn eine der vier Gruppen zeichnet als Versuchsaufbau Magen, Magensäure und Speiseröhre.

Bei Aufgabe II.1 fällt auf, dass die Schülerinnen und Schüler sehr intensiv den Kurvenverlauf diskutieren, dann aber in der eigentlichen Beantwortung der Aufgabe keinen direkten Bezug zur Kurve herstellen. Einer der vier Gruppen ist nicht klar, was in der Aufgabe von ihnen gefordert wird. Auch die Formulierung des Aufgabentextes von Aufgabe III.1. verursacht bei einigen Schülerinnen und Schülern Verständnisschwierigkeiten.

Fragestellung 1.2:

Lassen sich Aufgabenformate ausmachen, die den Lernenden besondere Probleme bereiten?

Auch zu besonderen Schwierigkeiten, die im Aufgabenformat begründet sind, liefert die Videoaufzeichnung Hinweise. In der Aufgabe I.1. geht aus dem Aufgabentext nicht hervor, ob eine oder mehrere Antwortalternativen angekreuzt werden sollen, zwei Gruppen fragen somit während der Bearbeitung nach. Das Format der Aufgabe III.5. stößt bei zwei der vier Gruppen auf Unverständnis, denn es wird ihnen nicht deutlich, welches Vorgehen in dieser Aufgabe von ihnen erwartet wird. Die letzte Aufgabe des Tests erfordert das Formulieren einer Antwort in Form eines Textes. Es fällt auf, dass die Schülerinnen und Schüler aus zwei Gruppen große Schwierigkeiten haben, ihre Lösung in Worte zu fassen, da ihnen z. T. die Begrifflichkeiten nicht klar sind (Aussage einer Schülerin: „Wie heißen die Dinger noch?“ „Die wandern da so durch“).

Fragestellung 2: Erkenntnisse aus der Itemanalyse

Die Fragebögen aus der Videoaufzeichnung fließen nicht in die statistische Auswertung ein, da die Aufgaben in Gruppen und in Diskussion bearbeitet wurden und daher die Schülerinnen und Schüler bessere Ergebnisse erzielen als die allein arbeitenden Probanden in Phase 2 der Pilotuntersuchung.

Aus der Videoaufzeichnung lassen sich jedoch erste Hinweise zu schwierigen bzw. leichten Aufgaben finden (Fragestellung 2.4). Bei der Bearbeitung der einzelnen Aufgaben durch die Gruppen zeigt sich, dass bei der Aufgabe I.1. Verunsicherung darüber auftritt, ob die Beantwortung wirklich so einfach ist (Aussage eines Schülers: „Aktivkohle. Alles andere ist zu leicht.“). Bei der Formulierung von Fragen (Aufgabe I.3.) fällt auf, dass es in den Gruppen besondere Probleme aufwirft, eine Frage an eine Biologielehrkraft zu stellen. Das Durchführen der Berechnung (Teilaufgabe I.5) bereitet allen Schülerinnen und Schülern dieser Klasse große Schwierigkeiten.

5.4.2 Zwischendiskussion 1

Aus den Ergebnissen der Gruppenbearbeitung der Testaufgaben lässt sich zusammenfassend festhalten, dass der Fragebogen insgesamt als Testinstrument geeignet ist. Bei einzelne Items müssen Schwierigkeit und Formulierung der Aufgabe überarbeitet werden, z. B. herrscht bei Aufgaben I.1 Unklarheit über die Anzahl der Antworten. Zudem lässt sich jetzt schon sagen, dass die Berechnungsaufgabe I.5 zu schwierig ist und vereinfacht werden muss. Der Begriff des „Modellversuchs“ in Aufgabe I.4 muss überdacht werden, zudem muss in Aufgabe II.1 deutlicher formuliert werden, dass die Grafik in die Beantwortung der Aufgabenstellung einbezogen werden soll. Auch die Formulierung des Aufgabentextes von Aufgabe III.1 und III.5 muss einer Überarbeitung unterzogen werden, zudem muss im ersten Fall eine kritische Überprüfung der Distraktoren vorgenommen und im letzten Fall zusätzlich zum Aufgabentext das Format überprüft werden. In Hinblick auf die Musterlösung im Erwartungshorizont für Aufgabe I.2 muss überlegt werden, wie mit sehr ähnlichen Begriffen bei der Bewertung umgegangen wird.

5.4.3 Phase 2 der Pilotuntersuchung: Fragebogeneinsatz

Fragestellung 1: Aufgabenverständnis der Schülerinnen und Schüler

Fragestellung 1.1

Sind die Aufgabeninhalte für Lernenden verständlich formuliert?

Bei Aufgaben bzw. für Teilantworten bei mehrstufigen Aufgaben, für die geringe Lösungshäufigkeiten festgestellt werden, werden bei einer inhaltlichen Analyse der Antworten Hinweise auf

Gründe für Schwierigkeiten gesucht. Die Lösungshäufigkeiten sind in Tabelle 17 dargestellt. In der Tabelle sind einzelne Lösungshäufigkeiten durch Fettdruck hervorgehoben, die sich dadurch auszeichnen, dass sie verhältnismäßig hoch bzw. niedrig sind.

Testaufgabe	Aufgabenanforderung, Antwortstufen	gültige Antworten	fehlende Antworten	Lösungshäufigkeiten in %
I.1.	Auswählen einer Substanz als Gegenmittel bei Sodbrennen (2)	117	3	korrekt: 38,5 teilrichtig: 54,1
I.2.	Nennen wichtiger Kriterien bei der Medikamententwicklung	112	8	korrekt: 50,9 teilrichtig: 22,3
I.3	a) Formulieren von adressatengerechten Fragen	120	0	korrekt: 65,0 teilrichtig: 28,3
	b) Formulieren von Suchbegriffen für eine Internetrecherche	118	2	korrekt: 78,0 teilrichtig: 17,8
I.4	Erstellen einer Versuchszeichnung	91	29	korrekt: 11,0 teilrichtig: 29,7
I.5	a) Aufstellen einer Reaktionsgleichung	77	43	42,9
	b) Durchführen einer stöchiometrischen Berechnung	21	99	korrekt: 14,3 teilrichtig: 0
II.1	Beschreiben eines Kurvenverlaufs	117	3	24,8
II.2	Identifizieren von fehlerhaften Informationen in der grafischen Auftragung	92	28	48,9
II.3	Beurteilen der Wirkmechanismen eines Zahnpflegekaugummi	120	0	24,2
II.4	Beurteilen der Aussagekraft einer vorgestellten Studie	120	0	48,3
III.1	Identifizieren einer Stoffaufzählung, die ausschließlich Stoffe nennt, die in Wasser gelöst basisch reagieren	101	19	42,6
III.2.	Identifizieren der pH-Kurve, die die beschriebene Titration korrekt wiedergibt	112	8	67,0
III.3.	Treffen einer begründeten Entscheidung für eine Vorgehensweise bei verschütteter Säure	112	8	korrekt: 28,6
				teilrichtig: 31,9
III.4	Auswählen eines geeigneten Indikators aus gegebener Liste	108	12	71,3
III.5.	Nennen von Eigenschaften saurer Lösungen mit geeigneten Demoversuchen	91	29	korrekt: 28,6
				teilrichtig: 31,9
III.6	Erklären des Phänomens der elektrischen Leitfähigkeit	79	41	korrekt: 6,3
				teilrichtig: 44,3

Tab. 17: Lösungshäufigkeiten der Testaufgaben der Pilotstudie; Werte, die besonders hohe oder niedrige Lösungshäufigkeiten aufweisen, sind durch Fettdruck hervorgehoben.

Das Formulieren von Suchbegriffen (I.3) fällt den Schülerinnen und Schülern leicht, denn 95,8 % der Probanden bearbeiten die Aufgaben so, dass sie mindestens eine teilrichtige Lösung produzieren, als vollständig korrekt werden 78 % der Lösungen klassifiziert.

Auf die Aufgabe I.4 und I.5 wird an späterer Stelle eingegangen, hier sei nur darauf verwiesen, dass es nur 11 % der Probanden gelingt einen Versuchsaufbau zu zeichnen, der entsprechend dem Erwartungshorizont als korrekt eingestuft wird und nur 14,3 % eine korrekte Berechnung durchführen. Die Aufgabe zur Beschreibung des Kurvenverlaufs (II.1) und die Beurteilung der verschiedenen Aussagen zur Wirkungsweise von Zahnpflegekaugummi bewältigen nur etwa ein Viertel der Probanden korrekt, die letzte Testaufgabe zur Erklärung der Leitfähigkeit (III.6) können nur 6,3 % der Schülerinnen und Schüler korrekt lösen.

Auf die Aufgabe I.4 (Erstellen einer Versuchszeichnung) wird im Zusammenhang mit Fragestellung 4.2 eingegangen und die komplexe Multiple-Choice-Aufgabe II.3 wird im Rahmen der Fragestellung 3.2 und 4.2 näher betrachtet, da in beiden Fällen keine inhaltlichen Probleme festzustellen sind, sondern vielmehr der Erwartungshorizont und die Punktevergabe überarbeitet werden müssen. Auch bei der Rechenaufgabe I.5 liefert die inhaltliche Analyse der Antworten keine weiterführenden Hinweise.

Nur etwa ein Viertel der Schülerinnen und Schüler bearbeiten die Aufgabe II.1 entsprechend den Erwartungen, wie sie im Erwartungshorizont zu Grunde gelegt sind. Es fällt auf, dass die Probanden in ihrer Erklärung keinen (ausreichenden) Bezug zur Grafik herstellen, sofern sie überhaupt in ihrer Antwort auf die Änderungen des pH-Wertes im Mund und damit auf den Kurvenverlauf eingehen. Ein großer Teil der Schülerinnen und Schüler formuliert in der Antwort keine Beschreibung des Verlaufs und damit der pH-Wert-Veränderung, sondern geht direkt zu einer Interpretation des Kurvenverlaufs über (S 28: „Laut des Herstellers wird durch das Kauen des Kaugummi die Säure neutralisiert. Durch das Kauen werden also die schädlichen Säure neutralisiert“; S 29: „Das Kaugummi neutralisiert Säuren im Mund und erhöht den pH-Wert, bis dieser nicht mehr im Problembereich, sondern im neutralen Bereich liegt“; S 65: „Nach dem Kauen des Kaugummi wird der pH-Wert im Mund wieder neutralisiert.“).

Für die Testaufgabe III.6, in der die Schüler aufgefordert sind, das Phänomen der Leitfähigkeit auf der Ebene der Teilchen zu erklären, lässt sich festhalten, dass die Schülerinnen und Schüler aus der Videoaufzeichnung Probleme beim Formulieren ihrer Überlegungen hatten. Auch in den Fragebögen lassen sich Hinweise darauf finden, dass u. a. der Umgang mit fachlichen Begriffen bei dieser Aufgabe eine besondere Schwierigkeit darstellt (S 27: „Damit Wasser leiten kann, müssen metallische Stoffe im Wasser enthalten sein, Salzsäure besitzt diese metallischen Stoffe“; S 76: „Das alles hängt vom Salzgehalt ab“; S 112: „Durch Säuren ist im Wasser etwas, was die Teilchen besser leitet“). Darüber hinaus zeigt sich bei dieser Aufgabe, dass der Begriff der Salzsäure die Probanden dazu verleitet, den Begriff des Salzes aufzugreifen (S 31: „Die Salzsäure enthält mehr Salze, die den elektrischen Strom besser leiten können. Im Wasser sind hingegen

kaum oder gar keine Salzmoleküle vorhanden“; S 40: „Alle Formen von Salzen, ob als Lösung oder Säure, leiten den elektrischen Strom“).

Fragestellung 1.2:

Lassen sich Aufgabenformate ausmachen, die den Lernenden besondere Probleme bereiten?

Die Auswertung der Fragebögen gibt keine weiteren Hinweise auf besondere Schwierigkeiten, die in dem jeweiligen Format der einzelnen Aufgabe begründet liegen, nur in Einzelfällen ist Aufgabe III.5 mit einem Fragezeichen versehen.

Fragestellung 2: Erkenntnisse aus der Itemanalyse

Fragestellung 2.1:

Reliabilität des Gesamttests: Welche Trennschärfen besitzen die einzelnen Items?

Eine Berechnung der Reliabilität des Gesamttests der Pilotstudie zeigt, dass der Test insgesamt bei Berücksichtigung aller 17 Testitems einen Wert von Cronbachs $\alpha = .696$ besitzt. In dieser Skala sind jedoch sieben Items enthalten, die in Bezug auf die gemessene Skala eine schlechte Trennschärfe besitzen. Wird die Reliabilität einer Skala bestimmt, die die sieben Items unberücksichtigt lässt, ergibt sich eine interne Konsistenz dieser Skala von Cronbachs $\alpha = .773$ (die Trennschärfen aller enthaltenen Items liegt über $.20$). Durch schlechte Trennschärfen zeichnen sich folgende Aufgaben aus: I.3 (Formulieren von a) Fragen und b) Suchbegriffen), I.4 (Erstellen der Versuchszeichnung), I.5 (Durchführen von Berechnungen), II.1 (Beschreiben des Kurvenverlaufs), II.3 (Beurteilung der Wirkungsweisen des Kaugummis) und II.4 (Beurteilen der Aussagekraft der Studie). Eine Zusammenstellung der entsprechenden Werte liefern die Übersichten in den Tabellen 40 und 41 im Anhang D dieser Arbeit (S. 181).

Fragestellung 2.2:

Lassen sich bei den Multiple-Choice-Aufgaben problematische Distraktoren finden, die die Probanden z. B. unbeabsichtigt „in die Irre führen“?

Im Test sind insgesamt sechs Multiple-Choice-Aufgaben enthalten, davon handelt es sich bei einer Aufgabe um ein komplexes Format dieses Aufgabentyps. Durch Häufigkeitsaussagen zum Ankreuzverhalten der Schülerinnen und Schüler in Hinblick auf die einzelnen Disktraktoren lassen sich Hinweise auf kritische Antwortalternativen erhalten, die die Aufgabenbearbeiter un-

beabsichtigt irreführen. Tabelle 42 im Anhang D dieser Arbeit (S. 182) zeigt die Ankreuzhäufigkeiten der jeweiligen Distraktoren der im Test eingesetzten Multiple-Choice-Aufgaben.

In der ersten Multiple-Choice-Aufgabe I.1 sind die Schülerinnen und Schüler dazu aufgefordert, aus den genannten Substanzen die auszuwählen, die sich bei Sodbrennen einsetzen lassen. Vollständig korrekt ist nur eine Antwort D („basisch wirkendes Salz“, angekreuzt von 57 % der Schülerinnen und Schüler²), aber aus rein chemischer Sicht ist auch die Lösung B („Laugen“, ebenfalls von 57 % gewählt) gerechtfertigt. Die beiden Distraktoren A („Säuren“) und E („Aktivkohle“) werden in zwei Fällen nur zu unter 2 % angekreuzt, für die Antwortalternative C („sauer wirkendes Salz“) entscheiden sich rund 5 % der Schülerinnen und Schüler.

In der komplexen Multiple-Choice-Aufgabe II.3 müssen die Probanden bei jeder Aussage eine „ja/nein“-Entscheidung treffen, wobei zwei Aussagen zutreffen und zwei Aussagen falsch sind. Die beiden korrekten Aussagen werden von 58 % („Durch Kauen angeregter Speichelfluss reinigt die Zähne.“) bzw. 59 % („Das Kauen eines Kaugummis reinigt die Zähne.“) der Schülerinnen und Schüler als richtige Aussagen erkannt, die Aussage „Zucker im Kaugummi konserviert die Zähne“ beurteilen 98 % der Schülerinnen und Schüler korrekt als falsch. Aber die letzte, inhaltlich falsche Aussage „Häufiges Kauen trainiert die Zähne“ halten immerhin ein Fünftel der Schüler für korrekt, jedoch wird in einigen Fragebögen die angegebene Aussage verbessert (der Begriff „Zähne“ wird gestrichen und durch „Kaumuskeln“ ersetzt) oder kommentiert mit „nicht die Zähne, sondern die Kaumuskeln“.

In der Multiple-Choice-Aufgabe II.4 sind die Schüler aufgefordert, über den Aussagegehalt der beschriebenen Studie zu entscheiden. Unter 7 % der Schüler sind der Ansicht, die Studie besage, Kaugummikauen sei gut für die Zähne, unter 2 % schließen sich der Aussage an, die Studie besage, Kaugummikauen sei schlecht für die Zähne. Etwas über die Hälfte der Probanden (54 %) entscheidet sich richtig für die Antwortalternative, dass die Studie gar nichts aussage, aber immerhin 45 % glauben, die Studie zeige, dass Kaugummikauen keine Wirkung auf Zähne habe³.

In Aufgabe III.1 sind die Schülerinnen und Schüler aufgefordert, die Aufzählung von Stoffen zu kennzeichnen, in der ausschließlich Stoffe genannt werden, die in Wasser gelöst basisch reagieren. Für die erste falsche Antwortalternative (NaOH, CH₃COOH, H₂SO₄) entscheiden sich etwa ein Fünftel der Probanden (22 %), für die zweite falsche Alternative (NaCl, NH₃, KOH) sogar über ein Viertel der Aufgabenbearbeiter (28 %), auch die letzte falsche Alternative (CH₃CH₂OH, CaO, HCl) halten immerhin noch 13 % der Schülerinnen und Schüler für richtig. Die richtige

² Prozentangabe bezieht sich im Folgenden immer nur auf die Probanden, die die Aufgabe überhaupt bearbeitet haben, nicht auf die Gesamtzahl der Probanden.

³ Die Summe der Prozentangaben ist > 100, da auch hier in einzelnen Fällen mehr als eine Alternative angekreuzt wurden.

Antwort (NaOH, CaO, NaHCO₃) wird von 47 % der Schülerinnen und Schüler als solche erkannt.

In der Aufgabe III.2 gilt es, die pH-Kurve auszuwählen, die die beschriebene Neutralisationstiteration richtig beschreibt. Den richtigen Kurvenverlauf kreuzen 67 % der Probanden an, die erste Alternative (Kurve endet bei pH = 7) erscheint 16 % der Lernenden als korrekter Kurvenverlauf, auf die übrigen drei Darstellungen entfallen 5 bzw. 8 %.

In der letzten Multiple-Choice-Aufgabe des Tests sollen die Schülerinnen und Schüler aufgrund angegebener pH-Bereiche entscheiden, mit welchem Indikator sie die Neutralität von aufgefangenem Regenwasser überprüfen würden. Für die richtige Antwortalternative (Bromthymolblau: pH 6,2 – 7,6) entscheiden sich 71 % der Probanden, immerhin 19 % wählen den zweiten Distraktor (Phenolphthalein: pH 8,0 – 9,8), für den letzten (Methylorange: pH 3,1 – 4,4) entscheiden sich 9 % der Schülerinnen und Schüler.

Fragestellung 2.3:

Gibt es unter den Testitems einzelne Aufgabe, die in besonderer Weise Probanden geschlechtsspezifisch benachteiligen?

Mit Hilfe von t-Test-Berechnungen für alle eingesetzten Testaufgaben und der Betrachtung der Ankreuzhäufigkeiten einzelner Distraktoren von Schülerinnen und Schülern bei Multiple-Choice-Items werden Hinweise darauf gesucht, ob der Test Aufgaben enthält, die geschlechtsspezifisch wirken. Es soll ermittelt werden, ob männliche bzw. weibliche Probanden besonders gut oder schlecht in der Lage sind, eine bestimmte Testaufgabe zu lösen, und ob bestimmte Distraktoren weibliche bzw. männliche Probanden besonders stark zu einem falschen Ankreuzverhalten verleiten.

Die Berechnungen des t-Tests ergeben, dass sich nur bei einer Aufgabe ein signifikanter Unterschied im Lösungsverhalten zwischen Mädchen und Jungen ergibt. Aufgabe I.3 zum Formulieren von Fragen, die sich an Personen mit unterschiedlichem Expertenwissen richten, schneiden Schülerinnen signifikant besser ab als Schüler ($t_{(117,89)} = -2,94$, $p \leq 0,01$, $d = 0,52$). Der Mittelwert beträgt bei den Schülerinnen in dieser Aufgabe 1,77 bei einer Standardabweichung von 0,428, während der Mittelwert der männlichen Probanden bei dieser Aufgabe bei 1,47 mit einer Standardabweichung von 0,689 liegt. Die Werte der t-Test-Berechnung können Tabelle 43 im Anhang D (S. 183) entnommen werden.

Zur genaueren Betrachtung dieses Items werden die Lösungshäufigkeiten von vollständig richtigen und teilrichtigen Lösungen untersucht. Für Aufgabe I.3 zum Formulieren von Fragen wird festgestellt, dass bei den weiblichen Probanden in allen Fällen mindestens eine teilrichtige Lö-

sung produziert wurde, bei den männlichen Probanden gab es bei dieser Aufgaben auch Antworten, die als falsch eingestuft werden (11 % der Schüler erreichen bei dieser Aufgabe keinen Punkt), zudem erreichen deutlich mehr Mädchen als Jungen die volle Punktzahl (77 % der Mädchen gegenüber 56 % der Jungen).

Im Ankreuzverhalten einzelner Distraktoren der im Test enthaltenen Multiple-Choice-Aufgaben zeigen sich keine signifikanten Unterschiede zwischen männlichen und weiblichen Probanden. Die Ergebnisse der t-Test-Berechnung zum unterschiedlichen Ankreuzverhalten von Schülerinnen und Schüler im Anhang D in Tabelle 44 (S. 184) dargestellt.

Fragestellung 2.4.

Wie sieht die Schwierigkeitsverteilung der Items im Gesamttest aus? Welche Aufgaben erweisen sich in einer ersten Sichtung als besonders schwierig, welche als zu leicht?

Um die Schwierigkeit der einzelnen Aufgaben des Tests zu bestimmen, werden Schwierigkeitsindices berechnet. Für dichotome Items berechnet sich die Itemschwierigkeit p nach $p_i = R_i/N_i$, die Schwierigkeit p einer Aufgabe i liefert bei dichotomer Auswertung folglich der Quotienten der „Richtiglöser“ (R) und der Gesamtzahl der gegebenen Antworten (N) an und entspricht damit der Lösungshäufigkeit des Items (Bortz & Döring 2003).

Für Aufgaben, die mehrstufige Antwortmöglichkeiten zulassen, wird die Itemschwierigkeit nach

$$p_i = \frac{\sum_{m=1}^n x_{im}}{k_i \cdot n}$$
 bestimmt, wobei die Summe der erreichten Punkte (x_i) durch die maximal erreichbare Punktschritte dieses Items dividiert wird. Hierbei ergibt sich die maximal erreichbare Punktschritte aus der maximalen Punktzahl k , die eine Person bei diesem Item erreichen kann, und der Anzahl der antwortenden Personen (n). In die Berechnung der Schwierigkeitsindices fließen nur die bearbeiteten Antworten ein, nicht bearbeitete Items bleiben unberücksichtigt. Aus dieser Definition ergeben sich für die Schwierigkeiten der Aufgaben Werte von 0 (schwerstes Item) bis 1 (leichtestes Item) (Bortz et al. 2003). Die durchschnittliche Itemschwierigkeit des untersuchten Tests beträgt $p_{ges} = 0,48$.

In der vorliegenden Arbeit werden Aufgaben, deren Schwierigkeitsindex zwischen 0,00 und 0,35 liegt, als schwierig betrachtet. Liegt der Index eines Items im Bereich von 0,36 bis 0,65, wird die Aufgabe als mittelschwer eingestuft, befindet er darüber, handelt es sich um ein leichtes Item. Einen Überblick über die Schwierigkeitsindices der 17 Testitems liefert Tabelle 18: Im Fragebogen der Pilotuntersuchung sind fünf Items als schwer einzustufen, sieben Items lassen sich dem mittelschweren Bereich zuordnen, als leicht können wiederum fünf Items bezeichnet werden.

Testaufgabe	Aufgabenanforderung, (Antwortstufen)	gültige Antworten	fehlende Antworten	M	SD	p
Aufgaben mit <i>niedrigem</i> Schwierigkeitsgrad						
I.1.	Auswählen einer Substanz als Gegenmittel bei Sodbrennen (2)	117	3	1.32	.60	.66
I.3	a) Formulieren von adressatengerechten Fragen (2)	120	0	1.58	.62	.79
	b) Formulieren von Suchbegriffen für eine Internetrecherche (2)	118	2	1.74	.53	.87
III.2.	Identifizieren der pH-Kurve, die die beschriebene Neutralisationstitation korrekt wiedergibt	112	8	.67	.47	.67
III.4	Auswählen eines geeigneten Indikators aus gegebener Liste	108	12	.71	.45	.71
Aufgaben mit <i>mittlerem</i> Schwierigkeitsgrad						
I.2.	Nennen wichtiger Kriterien bei der Medikamententwicklung (2)	112	8	1.24	.85	.62
I.5	a) Aufstellen einer Reaktionsgleichung	77	43	.43	.50	.43
II.2	Identifizieren von fehlerhaften Informationen in der grafischen Auftragsung	92	28	.49	.50	.49
II.4	Beurteilen der Aussagekraft einer vorgestellten Studie	120	0	.48	.50	.48
III.1	Identifizieren einer Stoffaufzählung, die ausschließlich Stoffe nennt, die in Wasser gelöst basisch reagieren	101	19	.43	.50	.43
III.3.	Treffen einer begründeten Entscheidung für eine Vorgehensweise bei verschütteter Säure (2)	112	8	.92	.70	.46
III.5.	Nennen von Eigenschaften saurer Lösungen mit geeigneten Demoversuchen (2)	91	29	.89	.82	.45
Aufgaben mit <i>hohem</i> Schwierigkeitsgrad						
I.4	Erstellen einer Versuchszeichnung (2)	91	29	.52	.69	.26
I.5	b) Durchführen einer stöchiometrischen Berechnung (2)	21	99	.29	.72	.14
II.1	Beschreiben eines Kurvenverlaufs	117	3	.25	.43	.25
II.3	Beurteilen der Wirkmechanismen eines Zahnpflegekaugummi	120	0	.24	.43	.24
III.6	Erklären des Phänomens der elektrischen Leitfähigkeit (2)	79	41	.57	.61	.28

Tab. 18: Schwierigkeitsindices p der einzelnen Testaufgaben. In der Tabelle gibt die Zahl hinter der jeweiligen Aufgabenanforderung die Anzahl der für das jeweilige Item zugelassenen Antwortalternativen an, darüber hinaus werden Mittelwert und Standardabweichung ausgewiesen.

Bei den mehrstufigen Items wird zusätzlich zur Itemschwierigkeit die Lösungshäufigkeit der einzelnen Antwortstufen betrachtet (dargestellt in Tab. 19), um Aussagen darüber treffen zu können, an welcher Stelle eine Teilantwort unsinnig erscheint. Besonders auffällige Ergebnisse werden in der Darstellung hervorgehoben, auf sie wird im Folgenden Bezug genommen.

Test- aufgaben	Antwortalternativen a) vollst. u. korrekt b) teilrichtig c) falsch	Anzahl Ant- worten		gültige Prozent
		Σ gesamt		
I.1.	a) Alternative „basisches Salz“ angekreuzt	45		38,5
	b) Alternative „basisches Salz“ und „Lauge“ oder nur Lauge“ angekreuzt	64		54,7
	c) falsche Alternative	8		6,8
		Σ 117		
I.2	a) nennt drei sinnvolle Kriterien	57		50,9
	b) nennt zwei sinnvolle Kriterien	25		22,3
	c) nennt weniger als zwei sinnvolle Kriterien	30		26,8
		Σ 112		
I.3	a) formuliert drei adressatengerechte Fragen	78		65,0
	b) formuliert zwei adressatengerechte Fragen	34		28,3
	c) formuliert nur eine adressatengerechte Frage	8		6,7
		Σ 120		
	a) formuliert drei geeignete Suchbegriffe	92		78,0
	b) formuliert zwei geeignete Suchbegriffe	21		17,8
	c) formuliert nur einen geeigneten Suchbegriff	5		4,2
		Σ 118		
I.4	a) Zeichnung gibt einen sinnvollen Aufbau vollständig wieder	10		11,0
	b) Zeichnung gibt einen sinnvollen Aufbau mit kleineren Schwächen wieder	27		29,7
	c) geschlossene Apparatur, zu unvollständig, falsch	54		59,3
		Σ 91		
I.5	a) vollständige und richtige Berechnung	3		14,3
	b) richtige Berechnung der molaren Massen von NaHCO ₃ und HCl	0		0
	c) falsche Rechnung	18		85,7
		Σ 21		
III.3	a) entscheidet sich mit richtiger Begründung	23		20,5
	b) entscheidet sich für Natronlauge, weil es neutralisiert	57		50,9
	c) begründet Entscheidung nicht oder falsch	32		28,6
		Σ 112		
III.5	a) nennt zwei Eigenschaften mit je einem Demoversuch	26		28,6
	b) nennt zwei Eigenschaften, mind. einer davon mit Demoversuch	29		31,9
	c) nennt Falsches oder weniger als unter b)	36		39,6
		Σ 91		
III.6	a) erklärt Leitfähigkeit, geht dabei auf saure Lösungen und dest. Wasser ein	5		6,3
	b) erklärt Leitfähigkeit generell, geht dabei nicht oder nur z. T. auf saure Lösung/dest. Wasser ein	35		44,3
	c) falsche Erklärung	39		49,4
		Σ 79		

Tab. 19: Lösungshäufigkeiten der einzelnen Antwortstufen bei mehrstufigen Items.

Die Vermutung, dass bei einigen Aufgaben eine vollständig korrekte Lösung sehr selten vorkommt, wird durch diese Betrachtung bestätigt: Bei Aufgabe I.4. (Erstellen einer Versuchszeichnung) gelingt es nur 11,0 % der Schülerinnen und Schüler eine korrekte Versuchszeichnung zu entwerfen, die den im Erwartungshorizont formuliertem Ansprüchen für eine volle Punktzahl entspricht. Immerhin 29,7 % der Probanden erstellen eine Zeichnung, die mit kleineren Schwächen als teilrichtig gewertet wird. Die Rechenaufgabe wird insgesamt von 21 Schülerinnen und

Schülern bearbeitet, von denen nur drei die vollständig richtige Lösung errechnen. Für eine teilrichtige Lösung war es notwendig, die molaren Massen zu berechnen, dieser Zwischenschritt gelingt keinem der restlichen 18 Schülerinnen und Schüler. Auch bei der Aufgabe zur Erklärung der elektrischen Leitfähigkeit erhalten nur wenige Probanden die volle Punktzahl (6,3 %), aber immerhin 44,3 % sind in der Lage, das Phänomen Leitfähigkeit generell zu erklären, d. h. ohne dabei speziell auf den in der Aufgabe erfragten Unterschied zwischen einer sauren Lösung und destilliertem Wasser einzugehen.

Fragestellung 2.5:

Sind die Testaufgaben geeignet, die Leistungen der Schülerinnen und Schüler hinreichend differenziert darzustellen?

Ein Test muss eine möglichst breite Streuung der Aufgabenschwierigkeiten aufweisen, damit Probanden mit unterschiedlichen Fähigkeiten annähernd gleich gut differenziert werden (Bortz et al. 2003). Demnach sind mittlere Itemschwierigkeiten zwischen 0,20 – 0,80 günstig (Durchschnittsschwierigkeit des eingesetzten Testinstruments: $p_{\text{ges}} = 0,48$). Bei Betrachtung der in Fragestellung 1.3. ermittelten Aufgabenschwierigkeiten (Tab. 18) unter diesem Gesichtspunkt stellt man fest, dass Aufgabe I.5 mit $p = 0,14$ zu schwierig und Aufgabe I.3 mit $p = 0,87$ zu einfach zu lösen sind und damit jenseits der empfohlenen Schwierigkeitsgrenzen liegen. Die restlichen 15 Items verteilen sich über die angegebene Schwierigkeitsspanne (vgl. Tab. 18).

Für die folgenden Betrachtungen zur differenzierten Darstellung der Leistungen der Schülerinnen und Schüler wird zunächst die Verteilung der erreichten Punktzahlen im Gesamttest in den Blick genommen. Insgesamt konnten im Test der Pilotstudie maximal 26 Punkte erreicht werden. In diesem Zusammenhang muss berücksichtigt werden, dass es sich bei der Gesamtpunktzahl um einen ungewichteten, additiven Index (Bortz et al. 2003) handelt, und zudem muss beachtet werden, dass ein Teil der Aufgaben dadurch aufgewertet wird, dass mehrstufige Antworten möglich sind. Das bedeutet z. B., dass mit einer vollständig korrekten Antwort zu einem zweistufigen Item zwei falsche Antworten bei dichotomen Items kompensiert werden können. Um dieses Problem zu umgehen, wird auch bei den mehrstufigen Antworten die vollständig richtige Antwort mit 1 als höchste Punktzahl gewertet, für teilrichtige Antworten gibt es 0,5 Punkte, da nur zweistufige Antwortformate zum Einsatz kommen. Nach dieser Berechnung können entsprechend der Aufgabenanzahl im Test als Gesamtpunktzahl maximal 17 Punkte erreicht werden.

Ergebnisse aller Probanden im Gesamttest

Die niedrigste Punktzahl, die von Probanden im Test der Pilotstudie erreicht wird ist 1,0 (ein Proband), der höchste erreichte Wert liegt bei 14,5 (zwei Probanden). Die Verteilung wird in Abbildung 5 veranschaulicht. Im Durchschnitt werden 7,28 Punkte bei einer Standardabweichung von 2,81 erreicht.

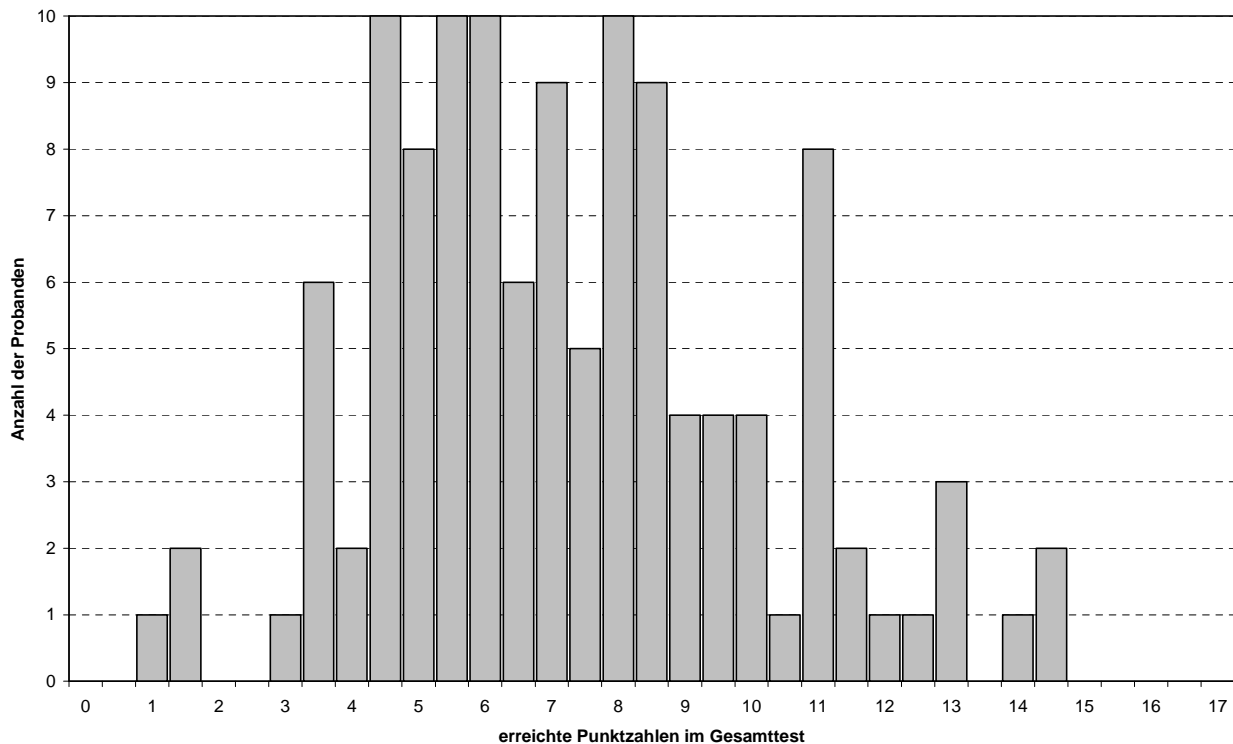


Abb. 5: Die Häufigkeitsverteilung der erreichten Gesamtpunktzahlen aller Probanden ($n = 120$); erreichte Gesamtpunktzahlen (max. Punktzahl 17) aller Probanden ($n = 120$); $M = 7,28$, $SD = 2,81$.

In der Betrachtung der Verteilung der Probanden anhand der erreichten Gesamtpunktzahlen im Test zeigt sich, dass alle Schülerinnen und Schüler in der Lage sind, mindestens eine Aufgaben erfolgreich zu bearbeiten, die höchste erreichte Gesamtpunktzahl liegt mit 14,5 Punkten unter der maximal erreichbaren Punktzahl von 17. Drei Probanden befinden sich im unteren Leistungsbereich (1 – 1,5 Punkte), drei schneiden deutlich besser als die gesamte Stichprobe (14 – 14,5 Punkte), die übrigen verteilen sich auf Punktzahlen von drei bis 13 Punkte.

Ergebnisse im Gesamttest der einzelnen Klassen

Die folgende Abbildung 6 stellt das Abschneiden der teilnehmenden Klassen im Gesamttest dar. Insgesamt zeigt ein recht homogenes Bild der Klassenleistungen (im Vergleich dazu Steinhoff 2004), aber eine einfaktorielle Varianzanalyse ergibt, dass es signifikante Unterschiede im Abschneiden der Klassen gibt, $F(5,114) = 24,12$, $p \leq 0,001$. Die Bestimmung von Scheffé-Kontrasten zeigt, dass Klasse 3 signifikant besser abschneidet als die übrigen fünf Klassen, der

Mittelwert der erreichten Punktzahl von Klasse 3 liegt bei 11,64 bei einer Standardabweichung von 1,86. Eine Tabelle mit den entsprechenden Kennwerten befindet sich im Anhang dieser Arbeit (Anhang D, Tab. 45, S. 185).

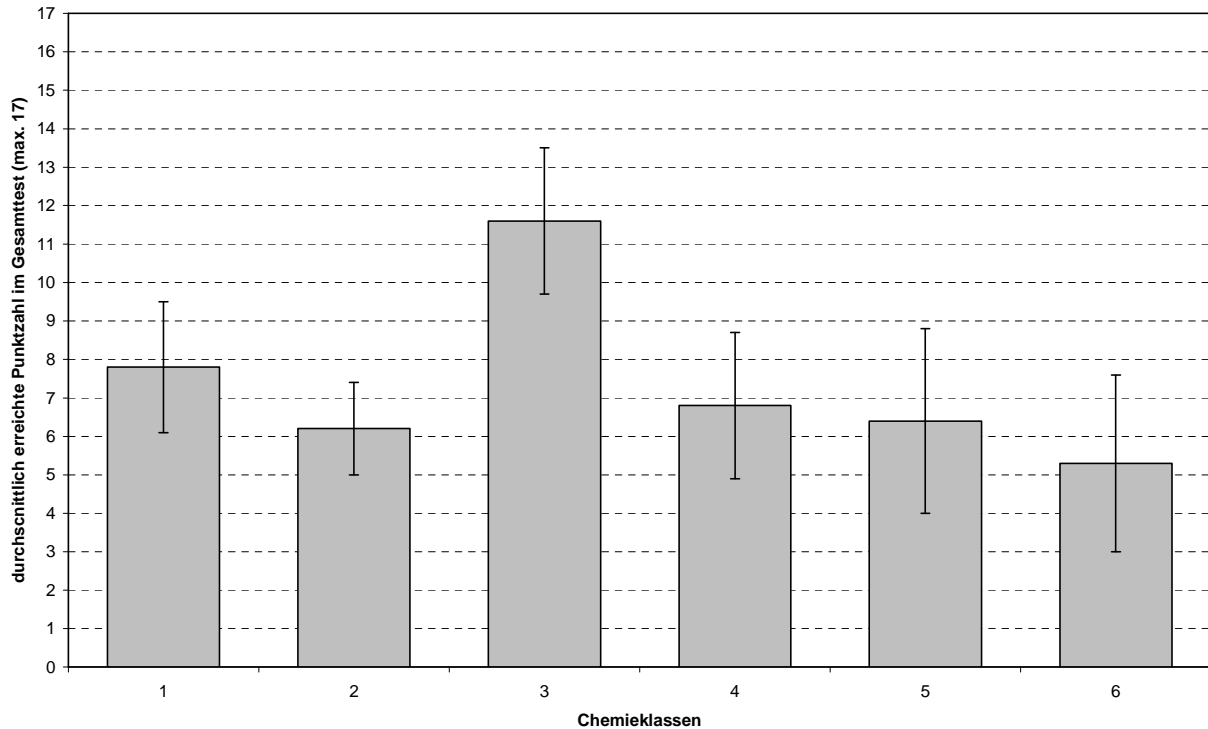


Abb. 6: Gesamtest: Leistung der teilnehmenden Klassen.

Ergebnisse der Probanden in Hinblick auf die ermittelte Fragebogendimension

Die Reliabilitätsanalyse (Fragestellung 2.1) hatte das Vorhandensein einer Dimension bestehend aus 10 Items ausgewiesen. Für diese Dimension wird im Folgenden die durchschnittlich erreichten Punkte in den sechs teilnehmenden Klassen im Vergleich aufgezeigt (1). Im Anschluss daran wird auf das unterschiedliche Abschneiden von männlichen und weiblichen Probanden in Bezug auf diese Fragebogendimension eingegangen (2).

(1) Vergleich über die teilnehmenden Klassen

Abbildung 7 stellt das Abschneiden der teilnehmenden sechs Klassen in Hinblick auf die zehn untersuchten Testaufgaben dar. Insgesamt zeigt sich wiederum ein homogenes Bild der Klassenleistungen, aber auch hier ergibt eine einfaktorielle Varianzanalyse, dass es auch bei Betrachtung der zehn Items Unterschiede im Abschneiden der Klassen gibt, $F(5,114) = 35,36$, $p \leq 0,001$. Die Bestimmung von Scheffé-Kontrasten liefert, dass Klasse 3 wiederum signifikant besser abschneidet als die übrigen fünf Klassen, der Mittelwert der erreichten Punktzahl von Klasse 3 liegt bei 8,69 bei einer Standardabweichung von 0,94 (weitere Kenndaten dargestellt in Tab. 20).

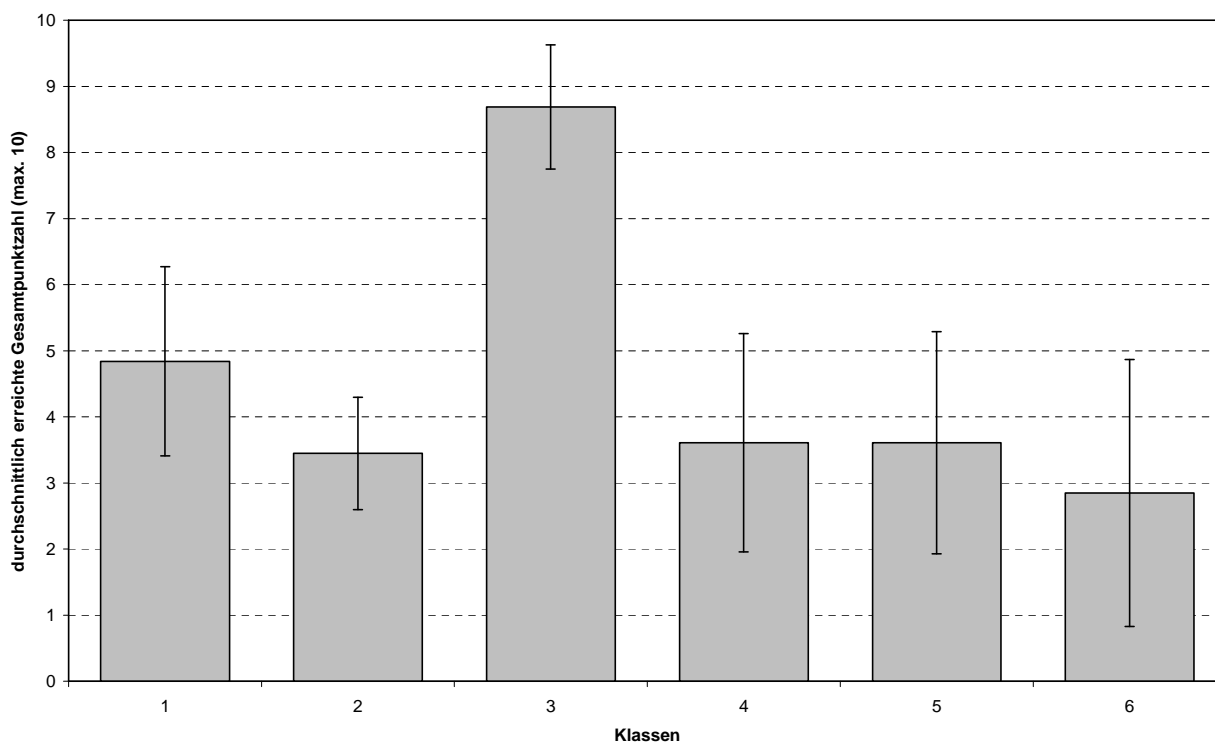


Abb. 7: Abschneiden der einzelnen Klassen hinsichtlich der ermittelten Fragebogendimension im Vergleich.

Klasse	N	M	SD	Min.	Max.
1	25	4,84	1,43	2,0	7,5
2	11	3,45	0,85	2,0	5,0
3	18	8,69	0,94	6,5	10,0
4	18	3,61	1,65	0,0	6,5
5	22	3,61	1,68	0,5	6,0
6	26	2,85	2,02	0,0	7,5
Σ 120		4,45	2,45	0,0	10,0

Tab. 20: Kenndaten zum Vergleich des Abschneidens der sechs Klassen hinsichtlich der zehn Items: Klassengröße, Mittelwerte, Standardabweichungen, Minimum und Maximum.

(2) Vergleich zwischen männlichen und weiblichen Probanden

Zwischen dem Abschneiden von männlichen und weiblichen Probanden gibt es keinen signifikanten Unterschied, die betrachtete Dimension des Fragebogens wird folglich nicht von Probanden unterschiedlicher Geschlechtszugehörigkeit verschieden gut bearbeitet. Männliche Probanden erlangen im Durchschnitt eine Punktzahl von 4,67 bei einer Standardabweichung von 2,43, von den weiblichen Probanden wird im Mittel eine Punktzahl von 4,11 bei einer Standardabweichung von 2,46 erreicht.

5.5 Diskussion

Aus den im vorhergehenden Abschnitt dargestellten Ergebnissen der Pilotuntersuchung ergeben sich verschiedene Veränderungen für die Aufgaben des Testinstruments der Hauptstudie, diese Änderungen werden im Folgenden aufgezeigt und begründet.

Darüber hinaus wird diskutiert, warum einzelne Aufgaben aus dem Testinstrument in der Hauptuntersuchung eingesetzt werden, obwohl sich in der Pilotuntersuchung gezeigt hat, dass sie in Bezug auf die gemessene Dimension eine schlechte Trennschärfe besitzen.

Zunächst werden in kurzer Form allgemeine Schlussfolgerungen der Ergebnisse mit ihren Folgen für die Gestaltung der Testaufgaben im Fragebogen aufgezeigt. Daran schließt sich eine detailliertere Betrachtung der einzelnen Fragestellungen an.

Zusammenfassend wird festgehalten, dass...

- sich das Testinstrument insgesamt als geeignet erweist, um die Leistungen von Lernenden am Ende der zehnten Klasse zum Thema „Säuren und Basen“ zu überprüfen.
- sich keine geschlechtsspezifische Benachteiligung im Test finden lässt.
- ein Großteil der Testaufgaben nicht grundlegend geändert werden muss.
- es anhand des Testinstruments möglich ist, Leistungsunterschiede zu erfassen.
- die Bearbeitungsdauer den Rahmen einer Schulstunde nicht übersteigt.
- der Fragebogen insgesamt eine mittlere Schwierigkeit besitzt.
- sich die Distraktoren insgesamt als brauchbar erweisen.
- sowohl der Erwartungshorizont als auch die Punktevergabe bei der Bewertung der Antworten der Schülerinnen und Schüler präzisiert und ausdifferenziert werden muss.

Im folgenden Abschnitt werden die Ergebnisse zu den einzelnen Fragestellungen aus Phase 2 der Pilotuntersuchung kurz zusammengefasst und im Anschluss in Hinblick auf mögliche Änderungen einzelner Testaufgaben und des gesamten Fragebogens diskutiert.

Abschließend wird in tabellarischer Form dargestellt, welche Änderungen die Testitems der Pilotstudie für den Einsatz in der Hauptstudie erfahren haben.

Fragestellung 1: Aufgabenverständnis der Schülerinnen und Schüler

In einer genaueren Betrachtung der Lösungshäufigkeiten für die (Teil-) Aufgaben zeigt sich, dass die Bewertung von einigen teilrichtigen Lösungen in Frage gestellt werden muss, weil die Aufgabe für die Probanden so leicht ist, dass ein großer Prozentsatz sie vollständig richtig löst (I.3, Formulieren von Suchbegriffen), weil die Aufgabe insgesamt so schwierig ist, dass niemand von den Probanden eine teilrichtige Lösung produziert (I.5, Berechnung) oder weil die Erwartungen an eine vollständig korrekte Lösung so hoch sind, dass der überwiegende Teil nur eine teilrichti-

ge Beantwortung der Aufgabe gibt (I.4, Versuchszeichnung, und III.6, Erklärung der Leitfähigkeit). Zudem fallen zwei Aufgaben (II.1, Kurvenbeschreibung, und II.3, Wirkungsweise von Zahnpflegekaugummi) durch verhältnismäßig geringe Lösungshäufigkeiten auf.

Die Aufgabe zur Nennung der Suchbegriffe bleibt mit mehrstufigem Bewertungsschema erhalten, die Sinnhaftigkeit der Kombination der drei Suchbegriffe wird aber in der Hauptuntersuchung bei der Bewertung der Aufgabenlösung berücksichtigt.

Für die Bewertung der Aufgaben I.4 wird für die Hauptuntersuchung ein dreistufiges Format der Bewertung ausgearbeitet, um den verhältnismäßig hohen Anteil der Schülerinnen und Schüler, die eine geschlossene Versuchsapparatur zeichnen, als eigene Gruppe erfassen zu können. Theoretisch können die Lernenden mit dieser Aufgabe bearbeiten, aber das gedankliche Durchspielen ihres entwickelten Versuchs, um ihn auf Praktikabilität zu überprüfen, unterbleibt in den meisten Fällen. Grund dafür könnte maßgeblich mangelnde Experimentiererfahrung sein.

Die Aufgabe I.5 wird stark vereinfacht, aber auch diese vereinfachte Version wird anhand eines zweistufigen Schemas bewertet. Die Anforderungen dieser Aufgaben waren zu hoch, denn Lernende können diese Aufgabe nicht lösen, selbst wenn Molmassenberechnungen bereits Unterrichtsthema waren.

Für die Aufgabe II.1, in der die Schülerinnen und Schüler aufgefordert sind, einen Kurvenverlauf zu beschreiben, zeigt sich, dass nur ein Viertel der Probanden die Aufgabe entsprechend der Musterlösung korrekt bearbeitet. Ein großer Teil der Probanden beschreibt nicht den Verlauf der Auftragung, sondern zieht Schlussfolgerungen aus der Grafik. Die Aufgabenstellung ist nicht ausreichend verständlich formuliert und muss dahingehend umformuliert werden, dass den Bearbeitenden verständlich wird, was in dieser Aufgabe von ihnen erwartet wird, nämlich zunächst eine Beschreibung des Kurvenverlaufs mit anschließender Interpretation. Man kann vermuten, dass Lernenden der Umgang mit grafisch repräsentierten Informationen (im Chemieunterricht) nicht hinreichend vertraut ist. Das korrekte Vorgehen bei der Auseinandersetzung mit einer Grafik ist auch in anderen Fächern von Bedeutung, aber entweder sind auch hier die Fähigkeiten wenig ausgeprägt oder die Lernenden übertragen ihre Kenntnisse z. B. aus dem Erdkundeunterricht nicht auf den Chemieunterricht.

Ebenfalls als unerwartet schwierig erweist sich Aufgabe II.3, in der die Probanden einzelne Aussagen zur Wirkungsweise von Zahnpflegekaugummi beurteilen sollen. Es wird davon ausgegangen, dass die Schwierigkeit der Aufgabe nicht in der Formulierung des Aufgabentextes oder dem Format der Aufgabe zu suchen sind. Für eine genauere, inhaltliche Betrachtung werden vielmehr die Antworthäufigkeiten für die vier Distraktoren genauer untersucht (Fragestellung 2.2).

Abschließend wird die Aufgabe III.6 betrachtet, denn nur ein geringer Prozentsatz von 6,3 % der bearbeitenden Schülerinnen und Schüler ist in der Lage eine Antwort zu formulieren, die den Anforderungen einer vollständig richtigen Lösung entspricht, wohingegen 44,3 % eine teilrichtige Antwort produzieren. Bei dieser Aufgabe zeigt sich bereits in der ersten Phase der Erprobung (Gruppenbearbeitung), dass die Probanden vor allem Probleme bei der Verbalisierung und dem Einsatz von fachlichen Begriffen haben. Einerseits wird die Formulierung des Aufgabentextes überarbeitet, um die Aufgabenanforderung verständlicher zu gestalten, andererseits wird die Musterlösung dahingehend verändert, dass die Anforderungen an die sprachliche Ausdrucksweise der Probanden gesenkt werden. Darüber hinaus wird davon abgesehen, ein konkretes Beispiel (Salzsäure) im Aufgabentext zu nennen, der Begriff wird durch die allgemeine Bezeichnung „saure Lösung“ ersetzt.

Die Aufgabenformate bereiten den Schülerinnen und Schülern keine auffälligen Schwierigkeiten, nur die tabellarische Gestaltung der Aufgabe III.5 wird überarbeitet, da sie in einigen Fällen mit einem Fragezeichen versehen wurde.

Fragestellung 2: Erkenntnisse aus der Itemanalyse

(1) Trennschärfe der Items

Die Ergebnisse zur Bestimmung der Reliabilität des Tests zeigen, dass einige Items enthalten sind, die eine ausgesprochen schlechte Trennschärfe in Hinblick auf die gemessene Fragebogen-dimension besitzen. Dennoch soll keines dieser sieben Items vollständig gestrichen werden, vielmehr wird bei zwei Aufgaben der Formulierung des Aufgabentextes verändert und in ihren Anforderungen überarbeitet, da es sich um Aufgaben handelt, die für die Probanden schwierig zu lösen sind (I.5b, II.1) bzw. unklare Anforderungen enthielten (II.1). Die beiden Aufgaben des zweiten Testteils (II.3, II.4) verbleiben wegen ihrer inhaltlichen Ausrichtung im Test, um sie in der Hauptuntersuchung gezielt qualitativ betrachten zu können. Während der Großteil der übrigen Aufgaben ein gewisses Grundverständnis von chemischen Sachverhalten erfordert, lassen sich diese Aufgaben mit „gesundem Menschenverstand“ bzw. Alltagserfahrungen oder Alltagsverständnis lösen, denn es geht um das Formulieren von adressatengerechten Fragen und Internet-Suchbegriffen, um die Beurteilung von Aussagen zur Wirkungsweise von Zahnpflegekaugummis und die Bewertung von Ergebnisaussagen einer durchgeführten Studie.

(2) Qualität der Distraktoren

Des Weiteren wurde untersucht, ob bei den verwendeten Aufgaben im Multiple-Choice-Format unbeabsichtigt irreführende Distraktoren eingesetzt sind. Es wird festgestellt, dass in der komplexen Multiple-Choice-Aufgabe II.3 immerhin etwa ein Viertel der Probanden die letztgenannte Aussage („Häufiges Kauen trainiert die Zähne.“) als korrekt beurteilen. Bei genauerer Betrachtung der Klassen zeigt sich, dass die Beurteilung dieser Aussage sehr unterschiedlich ausfällt, während nur 8,7 % der Probanden aus Klasse 1 die Aussage als richtig kennzeichnen, schließen sich aus Klasse 5 immerhin 31,8 % der Schülerinnen und Schüler dieser Aussage an (weitere Werte siehe Tab. 46, Anhang D, S. 185). Der Distraktor wird unverändert in den Haupttest übernommen, um zu untersuchen, ob sich das Vorhandensein dieser Alltagsvorstellung auch für die Stichprobe bestätigen lässt oder ob es sich um einen Klasseneffekt handelt.

Mit ähnlicher Begründung werden die Distraktoren der Aufgabe II.4 übernommen. Etwas weniger als die Hälfte der Probanden sind der Meinung, anhand der beschriebenen Studie eine Aussage treffen zu können. Im Gegensatz zur Aufgabe II.3 lässt sich hier kein so deutlicher Klasseneffekt ausmachen (siehe Tab. 47, Anhang D, S. 185), diese falsche Schlussfolgerung zieht sich gleichermaßen stark durch alle Klassen.

Bei der Aufgabe III.1 verteilen sich die nicht richtigen Antworten der Probanden über alle drei falschen Antwortalternativen: Der geringste Teil mit 13 % entfällt auf den letzten Distraktor, der in seiner Aufzählung HCl enthält, was vielen Schülerinnen und Schülern bekannt ist. Für die Aufzählung „NaOH, CH₃COOH, H₂SO₄“ entscheiden sich 22 %, weil gerade die Hydroxylgruppe des Essigsäuremoleküls durch Analogiebildung zur NaOH irreführend sein kann und zudem Schwefelsäure in der Formelschreibweise nicht ausreichend bekannt ist. Die zahlreichsten falschen Antworten (28 %) finden sich bei der Aufzählung „NaCl, NH₃, KOH“, während etwas weniger als die Hälfte der Probanden die richtige Aufzählung aus „NaOH, CaO, NaHCO₃“ als solche erkennt. Um den Schülerinnen und Schülern das Vorgehen nach dem Ausschlussprinzip deutlicher zu machen, wird die Aufgabenstellung umformuliert, damit deutlich wird, dass überlegt werden muss, was in Hinblick auf den pH-Wert passiert, wenn die genannten Stoffe in Wasser gelöst werden. Dazu ist es nicht erforderlich, dass ihnen jeder der genannten Stoffe bekannt ist, denn sobald einer als sauer (H₂SO₄, HCl) oder neutral (NaCl) reagierend identifiziert ist, kann eine Antwortalternative als falsch erkannt werden.

Bei der Multiple-Choice-Aufgabe III.2 wählen die Schülerinnen und Schüler aus verschiedenen pH-Wert-Auftragungen die Darstellung aus, die den beschriebenen Titrationsverlauf richtig wiedergibt. Ein verhältnismäßig hoher Prozentsatz von 16 % entscheidet sich für einen Verlauf von pH = 0 bis pH = 7, wobei sich wiederum deutliche Unterschiede in den einzelnen Klassen fest-

stellen lassen: drei Klassen kreuzen den Distraktor zu 0 bzw. 6,3 %, wohingegen sich aus den restlichen Klassen 36,4 %, 25 % bzw. 34,6 % der Schülerinnen und Schüler für diese Alternative entscheiden. In diesen drei letztgenannten Klassen zeigt sich, dass mindestens ein Viertel der Schülerinnen und Schüler ein falsches Verständnis von pH-Skala haben. Die Entstehung dieser fehlerhaften Auffassung kann nur mit Blick auf den tatsächlichen Unterricht geklärt werden. In der Aufgabe III.4 zur Auswahl eines geeigneten Indikators zeigt sich ebenfalls ein deutlicher Klasseneffekt, da aus Klasse 6 über die Hälfte der Probanden (57,1 %) für Phenolphthalein als geeigneten Indikator entscheiden (siehe Tab. 48, Anhang D, S. 186). Hier zeigt sich wiederum, dass entweder das Verständnis der pH-Skala unzureichend ist oder aber dass die Lernenden dieser Klasse nur diesen Säure-Base-Indikator kennen gelernt haben und ihn daher wählen.

(3) Geschlechterspezifität einzelner Items

Die Untersuchung zur Geschlechterspezifität einzelner Testaufgaben oder Distraktoren bei Multiple-Choice-Fragen ergibt, dass sich keine ausgesprochene Benachteiligung von Schülerinnen oder Schülern ergibt. Während die Distraktoren in Hinblick auf ihre geschlechtsspezifische Wirkung alle unproblematisch sind, findet sich unter den insgesamt 17 im Test eingesetzten Aufgaben eine, die von den Schülerinnen signifikant erfolgreicher bearbeitet wird als von den Schülern der Stichprobe. Die weiblichen Probanden bearbeiten Aufgaben I.3a, in der Fragen zu Medikamenten gegen Sodbrennen formuliert werden müssen, die sich an Personen mit unterschiedlicher Expertise richten, signifikant besser als ihre männlichen Mitschüler. Es wird vermutet, dass dieser Unterschied mit einer insgesamt besseren sprachlichen Ausdrucksfähigkeit der Schülerinnen im Vergleich zu den Schülern zusammenhängt. Sowohl wegen der Anzahl der Items als auch der Größe der Stichprobe könnte die Signifikanz eines einzigen Items in Hinblick auf ihre geschlechtsspezifische Wirkung aber auch in den Bereich der Zufälligkeit verwiesen werden.

(4) Schwierigkeitsverteilung

Für die Durchschnittsschwierigkeit des Gesamttests wird festgestellt, dass sie mit $p_{\text{ges}} = 0,48$ im gewünschten mittleren Schwierigkeitsbereich liegt, insgesamt ist der Anteil der leichten und schwierigen Aufgaben im Vergleich zu Aufgaben mit mittlerem Schwierigkeitsgrad verhältnismäßig hoch. Die als schwierig eingestuft Aufgaben sind bereits in anderen Zusammenhängen aufgefallen (u. a. Probleme im Aufgabenverständnis) und erfahren daher Veränderungen in den Aufgabenanforderungen, indem der Erwartungshorizont überarbeitet (I.4, III.6) eine (weitere) teilrichtige Lösung zugelassen (I.4, II.3) und / oder in der Aufgabenanforderung vereinfacht (I.5b) bzw. Aufgabenstellung präzisiert (II.1) wird.

Bei den als leicht eingestuften Aufgaben I.1, III.2 und III.4 handelt es sich in allen Fällen um Multiple-Choice-Aufgaben. Da sie sich im unteren Bereich der leichten Aufgaben befinden (die p-Werte für die Schwierigkeit liegen zwischen 0,66 und 0,71) werden sie unverändert in den Fragebogen der Hauptuntersuchung übernommen, nur die Reihenfolge der Indikatorenauflistung in Aufgaben III.4 wird verändert, damit die richtige Lösung (Bromthymolblau) nicht als erster Distraktor genannt wird. Die Formulierung von Internet-Suchbegriffen fällt den Probanden sehr leicht, daher wird in der Auswertung des Haupttests auch eine sinnvolle Kombination der drei Begriffe berücksichtigt.

Ebenfalls wenig Schwierigkeiten bereitet den Schülerinnen und Schülern das Formulieren von adressatengerechten Fragen, diese Aufgabe wird dennoch unverändert übernommen, da sich bereits in der Pilotuntersuchung zeigt, dass die Schwierigkeit davon abhängt, an welche Personengruppe die Frage gerichtet wird, so bereitet es den Probanden größere Schwierigkeiten, eine Frage an eine Biologielehrkraft zu richten als an eine Chemielehrkraft oder einen Apotheker.

(5) Erwartungshorizont

Für die Aufgaben I.4, I.5 und III.6 müssen die im Erwartungshorizont formulierten Ansprüche an eine korrekte Lösung überarbeitet werden, da sonst eine Differenzierung nicht möglich ist, weil ein Großteil der Lösungen als teilrichtig gewertet werden muss. Bei den Aufgaben I.3b und I.5 muss überlegt werden, ob eine Stufung beibehalten werden soll, weil die Probanden nur zu einem geringen Prozentsatz teilrichtige Lösungen produzieren. Da es aber generell sinnvoll erscheint, auch teilrichtige Lösungen zu erfassen und diese nicht als falsch zu werten, um eine differenziertere Auswertung zu ermöglichen, wird die Stufung von Antworten in allen Fällen beibehalten und zudem bei einigen Aufgaben zusätzlich eingeführt (dazu genauer Tab. 21, S. 103).

(6) Differenzierung von Schülerleistungen

Abschließend wurde der Fragebogen dahingehend untersucht, ob und inwieweit es damit möglich ist, die Leistungen der Lernenden differenziert darzustellen. Die Ergebnisse zeigen, dass es anhand der Testaufgaben möglich ist, sowohl gute als auch schlechte Testergebnisse zu erlangen. Die durchschnittlich erreichte Punktzahl liegt annähernd bei der mittleren Testpunktzahl, die Streuung ist entsprechend verhältnismäßig groß, was wiederum auf eine Verteilung der erreichten Punktwerte hinweist.

Betrachtet man die Leistungsverteilung der teilnehmenden Klassen im Gesamttest, bezogen auf die ermittelte Fragebogendimension bestehend aus zehn Items zeigen sich deutliche Unterschiede in den Klassen. Insbesondere Klasse 3 fällt im Vergleich zu den übrigen fünf Klassen durch

ein signifikant besseres Abschneiden im Gesamtttest und der ermittelten Fragebogendimension auf. Die Differenzierung der Leistungen der Probanden gelingt vor allem durch Items, die zur erfolgreichen Bearbeitung ein chemisches Wissen voraussetzen, während Items, die anhand von Alltagswissen und -erfahrungen korrekt bearbeitet werden können, weniger dazu geeignet sind, die Leistungen der Schülerinnen und Schüler zu differenzieren. Es wird vermutet, dass sich dieses Ergebnis durch einen Klasseneffekt erklären lässt. Während die Alltagskenntnisse bei allen Klassen in ähnlichem Maße vorhanden sind und als eher unabhängig von der unterrichtenden Lehrkraft angesehen werden können, hängen die Kenntnisse im Bereich „Schulwissen“ ganz entscheidend auch von der unterrichtenden Lehrkraft ab. Genauere Analysen zu Klasseneffekten werden in der Hauptuntersuchung möglich sein, da zusätzliche Informationen zum Chemieunterricht in der Auswertung zur Verfügung stehen.

Dieses Ergebnis deckt sich mit den ermittelten Trennschärfen der Testaufgaben, die Untersuchung zeigt, dass sich alle Items des letzten Testteils zu der ermittelten Fragebogendimension zählen. Darüber hinaus setzen auch die Aufgaben I.1 (Substanzen zur Behandlung von Übersäuerung), I.5 (Reaktionsgleichung) und II.2 (Fehler in der Grafik) zur Bearbeitung Wissen voraus, welches im Chemieunterricht vermittelt wird. Aufgabe I.2 (Kriterien zur Medikamentenentwicklung) zeichnet sich ebenfalls durch eine gute Trennschärfe aus, obwohl dieses Item auf der Basis von Alltagswissen beantwortet werden kann. In der Auswertung der Hauptstudie werden die von den Schülerinnen und Schülern genannten Kriterien inhaltlich analysiert, um zu untersuchen, ob die genannten Kriterien sich z. B. eher durch einen Bezug zum Chemieunterricht auszeichnen (durch Nennung von Begriffen z. B. zu „günstiger pH-Wert“, „chemische Verbindungen“ oder „Stoffmengen“) oder sich auf der alltagsnahen Ebene bewegen (wie z. B. Begriffe zu „Werbung“, Design der Verpackung“ oder „Preis“).

Für die übrigen Aufgaben ist ein Verständnis von Chemie nicht wichtig, so dass sie in Hinblick auf die ermittelte Fragebogendimension „Chemieunterricht“ nicht trennscharf sind. Zudem muss festgestellt werden, dass die restlichen sieben Items keine gemeinsame Skala bilden, da sie zwar alle nicht „Wissen aus dem Chemieunterricht“ messen, aber dennoch nicht ein gemeinsames Konstrukt „Alltagswissen“ abbilden, sondern scheinbar ganz unterschiedliche Facetten wiedergeben.

(7) Veränderungen im Testinstrument

Tabelle 21 verdeutlicht die Veränderungen, die an den Testaufgaben aufgrund der Ergebnisse der Pilotuntersuchung vorgenommen werden.

Testaufgabe im Pilottest	Vorgenommene Änderungen	Testaufgabe im Haupttest
I.1. (Gegenmittel Sodbrennen)	Keine Veränderungen	I.1. (Gegenmittel Sodbrennen)
I.2. (Kriterien Medikamententwicklung)	Keine Veränderungen	I.2. (Kriterien Medikamententwicklung)
I.3. (Fragenformulierung)	Keine Veränderungen	I.3. (Fragenformulierung)
(Suchbegriffe formulieren)	Keine Veränderungen	(Suchbegriffe formulieren)
I.4. (Versuchszeichnung erstellen)	In der eigentlichen Aufgabe keine Veränderungen, die Antworten werden dreistufig kategorisiert, der Erwartungshorizont entsprechend ausdifferenziert	I.4 (Versuchszeichnung erstellen)
I.5. (Reaktionsgleichung aufstellen)	Keine Veränderungen	I.5 (Reaktionsgleichung aufstellen)
(Berechnung durchführen)	Die Aufgabe wird vereinfacht, es soll nicht mehr die „Neutralisationsleistung“ berechnet werden, sondern es geht nun um die reine Berechnung der enthaltenen Stoffmenge. Das Antwortformat bleibt zweistufig (als teilrichtige Antwort wird die Berechnung der Molmasse von NaHCO_3 gewertet).	(Berechnung durchführen)
—	Es wird eine weitere Aufgabe zugefügt, die sich inhaltlich aus dem gestrichenen Teil der Rechenaufgabe ergibt, die Schüler werden dazu aufgefordert zu überlegen, was sie in Erfahrung bringen müssen (die überschüssige Säuremenge im Magen), um zu bestimmen, wie viele Tabletten eingenommen werden sollten.	(Informationen zur Entscheidungshilfe suchen)
II.1 (Kurvenverlauf beschreiben)	Zeichnung wird überarbeitet, Aufgabentext wird präziser formuliert, Aufgabe wird in zwei Teilaufgaben zerlegt (Beschreibung und Interpretation werden durch die Aufgabenformulierung eindeutig getrennt).	II.1 (Kurvenverlauf beschreiben)
—	—	(Kurvenverlauf interpretieren)
II.2 (Fehler identifizieren)	Keine Veränderungen	II.2 (Fehler identifizieren)
II.3 (Wirkungsweise von Zahnpflegekaugummi)	Die Aufgabe selbst erfährt keine Veränderung, es wird lediglich die Möglichkeit einer teilrichtigen Antwort eingeräumt.	II.3 (Wirkungsweise von Zahnpflegekaugummi)
II.4 (Studie beurteilen)	Keine Veränderungen	II.4 (Studie beurteilen)
III.1 (basisch reagierende Stoffe identifizieren)	Keine Veränderungen	III.1 (basisch reagierende Stoffe identifizieren)
III.2. (pH-Kurve identifizieren)	Keine Veränderungen	III.2. (pH-Kurve identifizieren)
III.3. (Vorgehensweise bei verschütteter Säure)	Keine Veränderungen	III.3. (Vorgehensweise bei verschütteter Säure)
III.4 (Indikator herausfinden)	Reihenfolge der Distraktoren wird verändert.	III.4 (Indikator herausfinden)
III.5. (Eigenschaften saurer Lösungen)	Inhaltlich bleibt die Aufgabe unverändert, der Aufgabentext und das Format erfahren eine Präzisierung, zudem wird die Aufgabe dreistufig ausgewertet.	III.5. (Eigenschaften saurer Lösungen)
III.6 (elektrische Leitfähigkeit erklären)	Aufgabentext wird umformuliert, unklare Formulierung herausgenommen, Salzsäure durch den Begriff „saure Lösung“ ersetzt.	III.6 (elektrische Leitfähigkeit erklären)

Tab. 21: Veränderungen in den Testitems.

5.6 Zusammenfassung

Auf der Basis der Ergebnisse sind Überarbeitungen der Testaufgaben in folgenden Bereichen vorgenommen worden: Aufgabeninhalte und Formulierungen des Aufgabentextes, Schwierigkeitsgrad, Formulierung des Erwartungshorizonts und Punktevergabe bei der Bewertung der Aufgaben. Alle Testaufgaben der Pilotstudie werden mit z. T. deutlichen Veränderungen in den Fragebogen der Hauptstudie übernommen. Zwei weitere Teilaufgaben kommen hinzu, indem Aufgaben aus der Pilotstudie differenzierter formuliert werden, um die Anforderung der Aufgabe zu vereinfachen bzw. stärker zu verdeutlichen.

Der Fragebogen ist ein Instrument, das dazu geeignet ist, erworbenes Wissen und erworbene Kompetenzen aus dem Chemieunterricht zu erfassen.

Die Ergebnisse der Pilotstudie zeigen erste Ansatzpunkte, an denen eine qualitative Analyse der Antworten der Schülerinnen und Schüler weitere Erkenntnisse liefern kann.

6 Hauptstudie: Einsatz von Testaufgaben zur Erfassung zentraler Kompetenzen im Chemieunterricht

Im folgenden Teil der Arbeit wird die Untersuchung der Hauptstudie vorgestellt, indem zunächst die Ziele und die daraus abgeleiteten Fragestellungen offen gelegt werden. Daran schließt sich die Beschreibung der Stichprobe und der eingesetzten Erhebungsinstrumente an. Es folgt die Darstellung der Ergebnisse zu den einzelnen Fragestellungen mit anschließender Diskussion.

6.1 Ziele der Untersuchung

In der Hauptstudie werden die überarbeiteten Testaufgaben der Pilotstudie als zentrales Untersuchungsinstrument eingesetzt, um Aussagen einerseits über die Qualität des Testinstruments, andererseits über das Abschneiden der teilnehmenden Lerngruppen treffen zu können.

Insgesamt werden Ergebnisse zu drei übergeordneten Fragestellungen berichtet: Der erste Fragenkomplex bezieht sich auf die statistische Prüfung des Fragebogens, anhand derer Aussagen zur Einsatzfähigkeit und Aussagekraft der überarbeiteten Testaufgaben getroffen werden, der zweite Komplex geht zunächst auf das Abschneiden einzelner Lerngruppen im Gesamttest ein, darüber hinaus werden die Probanden auf ihr Antwortverhalten hin analysiert. Im Rahmen der letzten Fragestellung wird auf einzelne Testitems, die von besonderem Interesse sind, eingegangen, wobei Lehrereinschätzungen, tatsächliche Lösungshäufigkeiten und inhaltliche Aspekte der Schülerantworten in Zusammenhang gebracht werden.

6.2 Fragestellungen

Aus den formulierten Zielen der Hauptstudie werden folgende Fragestellungen abgeleitet:

Fragestellung 1

Statistische Überprüfung des Testinstruments: Lassen sich mit dem Fragebogen verschiedene, für das naturwissenschaftliche Arbeiten zentrale Kompetenzen erfassen?

Fragestellung 2

Auswertung der Testergebnisse der Probanden: Welche Aussagen lassen sich zum Abschneiden einzelner Gruppen bzw. Klassen anhand des Tests treffen?

Fragestellung 3

Einzelbetrachtungen ausgewählter Testaufgaben: Welche Zusammenhänge lassen sich zwischen Lehrereinschätzungen, tatsächlichen Lösungshäufigkeiten und inhaltlicher Analyse der Schülerantworten ausgewählter Aufgaben herstellen?

Im Folgenden werden die übergeordneten Fragestellungen ausdifferenziert, indem untergeordnete Fragestellungen entwickelt und erläutert werden.

6.2.1 Fragestellung 1: Statistische Prüfung des Testinstruments

Ebenso wie der Test der Pilotstudie wird auch das Instrument der Hauptuntersuchung anhand von statistischen Kennwerten überprüft. Im Rahmen dieser Fragestellung wird zunächst mittels einer Reliabilitätsanalyse die interne Konsistenz des Fragebogens bestimmt. Darüber hinaus werden anhand einer Schwierigkeitsanalyse (vgl. Kap. 5.4.3) Aussagen über die Schwierigkeiten der einzelnen Items und über die Schwierigkeitsverteilung des gesamten Tests getroffen. Abschließend wird mittels einer Faktorenanalyse untersucht, ob der Test geeignet ist, verschiedene Kompetenzen abzubilden.

Fragestellung 1.1

Reliabilitätsanalyse: Mit welchen Reliabilitätswerte erfasst das Untersuchungsinstrument die Leistungen der Schülerinnen und Schüler?

Fragestellung 1.2

Faktorenanalyse: Ist es anhand des Tests möglich, entsprechend der zu Grunde gelegten Intention der Erfassung verschiedener Aspekte des naturwissenschaftlichen Arbeitens verschiedene Skalen abzubilden?

Fragestellung 1.3

Schwierigkeitsanalyse: Durch welche Schwierigkeitsindices sind die Testaufgaben gekennzeichnet?

6.2.2 Fragestellung 2: Auswertung der Leistungen der Schülerinnen und Schüler

Durch die Bearbeitung dieser Fragestellung werden die Leistungen der Lernenden in diesem Test deutlich. Es wird untersucht, ob sich Unterschiede zwischen den teilnehmenden Klassen und den unterschiedlichen Gruppen, die sich durch ihre Erfahrung mit der Konzeption Chemie im Kontext konstituieren, finden lassen. Eine weitere qualitative Auswertungsmethode kommt bei der Untersuchung des unterschiedlichen Antwortverhaltens der Probanden zum Tragen: Schülerinnen und Schüler lassen sich aufgrund ihres Antwortverhaltens in Hinblick auf die Bearbeitung ausgewählter Aufgaben zu Gruppen (Clustern) zusammenführen, diese Schülergruppen zeichnen sich durch gleiches Antwortverhalten hinsichtlich einer bestimmten Aufgabenauswahl aus. Es wird anhand dieser Fragestellung untersucht, welche Gruppen von Schülerinnen und Schülern sich bei den Probanden dieses Tests über *Chemie-im-Kontext*-Erfahrung und Schulklassenzuge-

hörigkeit hinaus bilden lassen und wie sich diese Gruppen hinsichtlich ihres Antwortverhaltens charakterisieren lassen.

Fragestellung 2.1

Gesamtleistungen: Welche Erkenntnisse liefert der Fragebogen in Bezug auf die unterschiedlichen Leistungen der Schülerinnen und Schüler?

Fragestellung 2.2

Clusterzentrenanalyse: Durch welche Merkmale lassen sich die Aufgabengruppen beschreiben, die bei verschiedenen Probandengruppen ein unterschiedliches Antwortverhalten hervorrufen?

6.2.3 Fragestellung 3: Einzelbetrachtung ausgewählter Aufgaben

Abschließend werden einzelne Items aus dem Gesamttest herausgegriffen und gesondert betrachtet. Zum einen werden in diesem Zusammenhang Aufgaben herausgegriffen, die sich in der Hauptstudie in Hinblick auf die untersuchte Dimension als wenig trennscharf erweisen. Abschließend finden sich in der gezielten Analyse von einzelne Aufgaben Items wieder, die von Gruppen mit und ohne *Chemie-im-Kontext*-Erfahrung (dabei wird nur ein Teil der Gesamtstichprobe betrachtet) unterschiedlich erfolgreich bearbeitet werden. In diesem Zusammenhang muss festgehalten werden, dass eine vergleichende Evaluation nicht möglich ist, da die Unterrichtsbedingungen nicht normiert sind. Darüber hinaus wurde die Stichprobengruppe mit Kontexterfahrung in der Unterrichtseinheit zum Thema Säuren und Basen nicht explizit nach *Chemie im Kontext* unterrichtet. Dennoch bietet das Thema Alltagsbezüge an, so dass es daher von Interesse ist, ob Schülerinnen und Schüler mit Kontexterfahrung anders abschneiden als Lernende, die bisher nicht nach der Konzeption unterrichtet wurden.

Fragestellung 3.1

Welche Besonderheiten kennzeichnen die Aufgaben, die bei der Betrachtung des Gesamttests aufgrund geringer Trennschärfen unberücksichtigt bleiben?

Fragestellung 3.2

Inwiefern zeigen sich Unterschiede in den Leistungen bei der Bearbeitung einzelner Aufgaben bei Betrachtung zweier Probandengruppen mit unterschiedlichem Erfahrungsstand hinsichtlich der Konzeption Chemie im Kontext bzw. einer bei klassenspezifischer Analyse?

6.3 Methode

Zunächst wird die Stichprobe beschrieben, die an der Hauptuntersuchung teilgenommen hat. Die Veränderungen des Testinstruments der Pilotphase zum Untersuchungsinstrument der Hauptstudie hinsichtlich Konstruktion der Testaufgaben, des Erwartungshorizonts und Bewertungsmusters wird in kurzer Form zusammengefasst. Zudem wird das zweite eingesetzte Instrument, der Fragebogen für die unterrichtenden Lehrkräfte, vorgestellt.

6.3.1 Beschreibung der Stichprobe

An der Hauptstudie nehmen insgesamt 359 Schülerinnen und Schüler aus 17 Schulklassen verschiedener Gymnasien aus fünf Bundesländern teil (Baden-Württemberg (eine Klasse), Bayern (acht Klassen), Hessen (fünf Klassen), Niedersachsen (zwei Klassen) und Schleswig-Holstein (eine Klasse)).

Alle Schüler wurden in der zehnten Klasse mit einer Unterrichtseinheit zum Thema Säuren und Basen unterrichtet. In einigen Fällen liegen zwischen Unterrichtseinheit und Befragung die Sommerferien, so dass ein Teil der befragten Schüler zum Zeitpunkt des Fragebogeneinsatzes bereits die 11. Klasse besucht und zwischen Unterrichtseinheit und Einsatz des Fragebogens demnach ein größerer Zeitraum von mindestens zwei Monaten liegt. Von den 17 teilnehmenden Klassen haben sechs Klassen (133 Probanden) keinerlei Erfahrung mit *Chemie im Kontext*, in vier weiteren Klassen ist das Thema Säuren und Basen als Einheit nach der Konzeption *Chemie im Kontext* unterrichtet worden, diese Einheit stellt für diese vier Klassen die erste *Chemie-im-Kontext*-Erfahrung dar. Sieben Klassen kennen die Konzeption bereits aus mehreren Einheiten. Die Auswahl der teilnehmenden Klassen fand nicht zufällig statt, vielmehr wurden gezielt Projektlehrkräfte angesprochen, so dass berücksichtigt werden muss, dass es sich bei den teilnehmenden Lerngruppen um Klassen handelt, die von z. T. in hohem Maße innovativen und engagierten Lehrkräften unterrichtet werden.

Eine Besonderheit stellen die bayerischen Klassen mit *Chemie-im-Kontext*-Erfahrung dar: In vier Klassen sind bisher fünf Kontexten unterrichtet worden, in einer weiteren bisher drei. Das Thema Säuren und Basen ist nicht als Kontexteinheit im Sinne der Konzeption unterrichtet worden, weil sich nach Aussage der unterrichtenden Lehrkräfte bei dieser Thematik auch ohne explizite Orientierung an *Chemie im Kontext* viele Alltagsbezüge herstellen lassen.

Die Dauer der Unterrichtseinheit variiert von 10 – 20 Stunden, zwei hessische Klassen fallen heraus, weil dort nur in drei bzw. vier Unterrichtsstunden grundlegende Definitionen des Themengebietes als Vorbereitung auf den Übergang zur Klasse 11 unterrichtet wurden. Das Durch-

schnittsalter der Probanden liegt bei 16,4 Jahren, der älteste Proband ist 20 Jahre, der jüngste 15 Jahre.

An der Hauptuntersuchung nahmen 150 weibliche und 209 männliche Testpersonen teil, d. h. ca. 41,8 % der Probanden sind weiblich, 58,2 % männlich. Fünf Lehrkräfte, die mit ihrer Klasse an der Untersuchung teilgenommen haben, besitzen keine Erfahrung mit der Unterrichtskonzeption *Chemie im Kontext*, die übrigen Lehrkräfte haben aus mindestens zwei, höchstens sieben Einheiten Erfahrung mit der Konzeption.

In Tabelle 22 ist im Überblick u. a. dargestellt, wie groß die einzelnen Klassen sind, ob sie *Chemie-im-Kontext*-Erfahrung haben, in welchem Kontext oder mit welchen Alltagsbezügen die Säure-Base-Thematik unterrichtet wurde.

Klasse (lfd. Nr.)	Klassen- größe	n (Mädchen (w)/Jungen (m))	Alters- durchschn. (in J.)	Ferien zw. Einheit u. Testeinsatz	Konzeptions- erfahrung d. Klasse	Name des Kontextes bzw. Alltagsbezüge in der Einheit (Dauer der Einheit in Schulstunden)
1	15	w= 11 m=4	16,4	nein	ja*	Sauer macht ... (10)
2	17	w= 11 m=6	16,4	nein	ja*	Alltagssäuren – alkalische Lösungen (16)
3	15	w= 8 m=7	16,4	nein	ja*	k. A. (14)
4	19	w= 9 m=10	16,1	nein	nein	Definitionen als Grundlage für den Übergang in Jg. 11 (4)
5	24	w= 7 m=17	17,4	nein	nein*	saurer Regen, Rosten, Kalken von Böden (k. A.)
6	17	w= 9 m=8	16,2	nein	ja*	Säuren – echt ätzend, Antacida (15)
7	25	w= 10 m=15	16,1	nein	ja*	Säuren – echt ätzend, Antacida (16)
8	23	w= 14 m=9	16,2	nein	nein	Definitionen als Grundlage für den Übergang in Jg. 11 (3)
9	21	w= 6 m=15	16,4	ja	nein	Reinigungsmittel, Konservierungsstoffe, Lebensmittel, Umwelt (10)
10	19	w= 6 m=13	16,5	ja	nein	k. A. (8)
11	27	w= 8 m=19	16,6	ja	ja*	Säuren und Basen im Haushalt (Haus- haltsreiniger, ...) (10) (anschließend „Kalk“-Kontext als Anwendung von Säuren und Basen)
12	25	w= 12 m=13	16,2	ja	ja*	Fruchtsäuren – Reinigungsmittel (10) (anschließend „Kalk“-Kontext als An- wendung von Säuren und Basen)
13	21	w= 5 m=16	16,5	ja	ja*	Reinigungsmittel, Backpulver, saurer Regen, Feuerlöscher (15) (anschließend „Kalk“-Kontext als Anwendung von Säuren und Basen)
14	24	w= 13 m=11	16,5	ja	ja*	Essig/Zitronensäure (Putzmittel), saurer Regen, technische Synthese von H ₂ SO ₄ , N-Kreislauf, ... (12) (anschließend z. T. „Kalk“-Kontext als Anwendung von Säuren und Basen)
15	19	w= 3 m=16	16,6	ja	ja*	Protonenübergang – Brönsted (10); (anschließend „Kalk“-Kontext als An- wendung von Säuren und Basen)

Klasse (lfd. Nr.)	Klassen- größe	n (Mädchen (w)/Jungen (m))	Alters- durchschn. (in J.)	Ferien zw. Einheit u. Testeinsatz	Konzeptions- erfahrung d. Klasse	Name des Kontextes bzw. Alltagsbezüge in der Einheit (Dauer der Einheit in Schulstunden)
16	27	w= 5 m=22	16,6	ja	nein	Reinigungsmittel, Biologie, Ernährung (20)
17	21	w= 13 m=8	15,4	nein	ja*	„Säurezirkel“ – Säuren im Haushalt: Reiniger, Reaktion mit unedlen Metallen, Naturstoffe als Indikator (Rotkohl, Tee), Mineralwasser, Knochen in Säuren (14)
	$\Sigma = 359$	$\Sigma_w = 150$ $\Sigma_m = 209$	$\bar{\varnothing} = 16,4$	$\Sigma_{\text{nein}} = 9$ $\Sigma_{\text{ja}} = 8$	$\Sigma_{\text{nein}} = 6$ $\Sigma_{\text{ja}} = 11$	

Tab. 22: Übersicht über die an der Hauptstudie teilnehmenden Klassen: Die mit * gekennzeichneten Angaben zur Konzeptionserfahrung der Lernenden kennzeichnen die vorhandene Konzeptionserfahrung der unterrichtenden Lehrkraft.

6.3.2. Erhebungsinstrumente der Hauptstudie

In der Hauptstudie wurden neben den Schülerinnen und Schülern zusätzlich die unterrichtenden Lehrkräfte befragt. In dieser Befragung ging es u. a. darum zu erfahren, welche Inhalte in den Unterrichtseinheiten zum Thema „Säuren und Basen“ aufgegriffen wurden, welche konkreten Inhalte und Anforderungen der Testaufgaben im Unterricht vermittelt wurden und wie die unterrichtenden Lehrkräfte die Fähigkeiten ihrer Klassen beim Bearbeiten der einzelnen Testaufgaben einschätzen.

(1) Erhebungsinstrument 1

Der Schülerfragebogen „*Sodbrennen, Kaugummi und Chemie*“ dient zur Erfassung von Kenntnissen und Kompetenzen der Schülerinnen und Schüler zum Konzept „Säuren, Basen und Neutralisation“. Die Veränderungen, die an den Testaufgaben anhand der Ergebnisse aus der Pilotstudie vorgenommen wurden, sind in Tabelle 21 (S. 103) bereits dargestellt worden, eine vollständige Version des Fragebogens findet sich im Anhang E (S. 187), der ausführliche Erwartungshorizont zu den Aufgaben im Anhang F (S. 196). Der Aufbau des Fragebogens insgesamt ist ausführlich im theoretischen Teil der vorliegenden Arbeit vorgestellt worden (Kap. 4.3) und wird daher an dieser Stelle nicht erneut aufgegriffen. Eine Kurzbeschreibung der einzelnen Items in gleicher Form, wie sie für die Aufgaben der Pilotstudie vorgenommen wurde, ist im Anhang der Arbeit (Anhang G, Tab. 50, S. 198) beigelegt. Aus dieser Übersicht gehen auch die im folgenden Text verwendeten Kurzbezeichnungen der Testaufgaben hervor.

Erwartungshorizont und Bewertung der Schülerantworten

Während in der Untersuchung der Pilotstudie nur zwei- und dreistufige Antwortformate verwendet werden, ist in der Untersuchung der Hauptstudie bei zwei Aufgaben eine weitere Differenzie-

rung vorgenommen, in zwei Fällen gibt es nun zusätzlich zwei Abstufungen für teilrichtige Lösungen, d. h. diese Aufgaben werden nach einem vierstufigen Antwortschema bewertet.

Tabelle 23 gibt Aufschluss darüber, nach welchem Antwortschema die einzelnen Testaufgaben bewertet werden.

Aufgabe	Antwortschema der Bewertung		
	zweistufig	dreistufig	vierstufig
I.1 „Gegenmittel Sodbrennen“		X	
I.2 „Aspekte der Medikamentenentwicklung“		X	
I.3a „Fragen formulieren“		X	
I.3b „Internet-Suchbegriffe“		X	
I.4 „Versuch zum Gas auffangen“			X
I.5a „Reaktionsgleichung“	X		
I.5b „Berechnung“		X	
I.5c „Informationen einholen“		X	
II.1a „Grafik beschreiben“	X		
II.1b „Grafik interpretieren“		X	
II.2 „Fehlersuche“	X		
II.3 „Kaugummiwirkung“		X	
II.4 „Studie bewerten“	X		
III.1 „basische Stoffe“	X		
III.2 „pH-Kurven“	X		
III.3 „Ratschlag“		X	
III.4 „Indikatoren“	X		
III.5 „Demonstrationsversuche“			X
III.6 „Leitfähigkeit“		X	

Tab. 23: Antwortschema der Bewertung für die Testaufgaben der Hauptuntersuchung.

Die Bewertung der Schülerantworten aus dem Fragebogen der Pilotstudie hat Hinweise darauf geliefert, bei welchen Aufgaben und für welche Aspekte der Erwartungshorizont differenzierter dargestellt werden muss, um eine Objektivität in der Auswertung gewährleisten zu können.

Objektivität der Bewertung der Aufgaben mit offenem Antwortformat

Für die Überprüfung der Objektivität der Bewertung der Aufgabenbeantwortungen, die von den Schülerinnen und Schülern frei formuliert werden, wird bei 40 % der erfassten Fragebögen (entspricht 144 Fragebögen) eine doppelte Codierung vorgenommen, in dem aus jedem Klassensatz 40 % von einer weiteren Person ausgewertet werden.

Da es sich mit elf Aufgaben um verhältnismäßig viele Items mit offenem Antwortformat handelt, werden die Aufgaben auf zwei Personen verteilt (zur Aufgabenverteilung: Tab. 24), bei den rest-

lichen in Tabelle 14 nicht aufgeführten Aufgaben handelt es sich um Aufgaben, deren Antwort im Multiple-Choice-Format, in Form einer Reaktionsgleichung oder einer Rechnung gegeben wird.

Codierte Aufgaben	
Interrater 1	I.2 (Aspekte der Medikamentenentwicklung)
	I.3.a (Fragen formulieren)
	I.3.b (Internet-Suchbegriffe)
	I.4 (Versuch zum Gasauffangen)
	I.5.c (Informationen einholen)
	III.3 (Ratschlag)
	III.5 (Demonstrationsversuche)
Interrater 2	II.1.a (Grafik beschreiben)
	II.1.b (Grafik interpretieren)
	II.2. (Fehlersuche)
	III.6 (Leitfähigkeit)

Tab. 24: Verteilung der Testaufgaben auf die Interrater 1 und 2.

Als Hauptkriterium wird das Übereinstimmungsmaß Cohens Kappa herangezogen, als zusätzliches Nebenkriterium die direkte Übereinstimmung genannt (Bortz et al. 2003).

In Tabelle 25 sind die Ergebnisse für die Aufgaben zusammengestellt, für die die Übereinstimmung in der Bewertung festgestellt wird: Aus den Daten ergibt sich mit einer Ausnahme für alle Aufgaben eine sehr gute Übereinstimmung in der Bewertung. Für Aufgabe II.1.a kann eine gute Übereinstimmung festgestellt werden.

Testaufgabe	Cohens Kappa (κ)	direkte Übereinstimmung in %	N
I.2	.87	92,1	140
I.3.a	.80	87,9	141
I.3.b	.85	93,1	144
I.4	.94	95,8	120
I.5.c	.93	97,4	119
II.1.a	.63	83,7	135
II.1.b	.71	82,9	140
II.2	.86	92,9	113
III.3	.83	90,2	132
III.5	.86	90,1	91
III.6	.78	85,6	97

Tab. 25: Bewertungsübereinstimmungen bei der Antwortcodierung der frei formulierten Schülerantworten.

(2) Erhebungsinstrument 2

Die unterrichtenden Lehrkräfte erhalten ebenfalls einen Fragebogen, der neben allgemeinen Fragen zu Dauer und Inhalten der Einheit zum Thema „Säuren, Basen und Neutralisation“ und den Erfahrungen mit der Unterrichtskonzeption *Chemie im Kontext* alle Testaufgaben des Schülerfragebogens in z. T. verkürzter Form enthält. Zu jeder dieser Testaufgaben werden die Lehrkräfte gebeten, ihre Einschätzung dazu abzugeben,

- ob nach ihrer Ansicht die notwendigen Kenntnisse und Fähigkeiten im Unterricht vermittelt wurden,
- welcher Prozentsatz ihrer Klasse die jeweilige Aufgabe vollständig und korrekt zu lösen vermag,
- wie viel Prozent ihrer Klasse Interesse an dieser Art der Aufgabe finden und
- ob diese Art der Aufgabengestaltung Bedeutung für ihren Unterricht hat.

Um vor allem eine Einschätzung zu den bereitgestellten Kenntnisse und Fähigkeiten und zu der Lösungshäufigkeit in ihrer Klasse vornehmen zu können, wird für jede Aufgabe kurz dargestellt, was bei einer vollständigen und korrekten Lösung erwartet wird und welche Kenntnisse und Fähigkeiten bei der Bearbeitung der jeweiligen Aufgabe zentral sind. Der Fragebogen für die Lehrkräfte findet sich im Anhang H dieser Arbeit (S. 199).

6.4 Ergebnisse

Im Folgenden werden die Ergebnisse der Untersuchung gegliedert nach den Fragestellungen aufgezeigt. Dazu wird zunächst in sehr kurzer Form auf die Ergebnisse des Lehrerfragebogen eingegangen, ein Rückgriff auf die Ergebnisse aus diesem Fragebogen erfolgt jeweils im Zusammenhang mit der Interpretation von Klassenergebnissen an entsprechender Stelle in der Ergebnisdarstellung zum Schülerfragebogen. Im Anschluss an die Ergebnisdarstellung des Fragebogens für die Lehrkräfte werden die Ergebnisse des Schülerfragebogens aufgezeigt.

6.4.1 Ergebnisse zur Auswertung des Lehrerfragebogens

Die folgenden Tabellen 26a und 26b zeigt die Einschätzung der Lehrkräfte für alle Aufgaben des Tests in Hinblick darauf, inwieweit die zur Bearbeitung der Aufgabe benötigten Kenntnisse und Fähigkeiten im vorangegangenen Unterricht bereitgestellt worden sind und in welchem Maße die jeweilige Aufgabenart generell Bedeutung für ihren Unterricht besitzen.

Aufgabe	N	fehlende Angaben	M	SD	eher ja (%)	eher nein (%)
Kenntnisse/Fähigkeiten im U. bereitgestellt						
I.1	17	1	0,81	0,403	81,2	18,8
I.2	17	1	0,00	0,000	0,0	100,0
I.3a	17	1	0,38	0,500	37,5	62,5
I.3b	17	1	0,38	0,500	37,5	62,5
I.4	17	1	0,75	0,447	75,0	25,0
I.5a	17	1	0,73	0,458	73,3	26,7
I.5b	17	1	0,63	0,500	62,5	37,5
I.5c	17	1	0,44	0,512	62,5	37,5
II.1a	17	1	0,63	0,500	43,8	56,2
II.1b	17	1	0,63	0,500	62,5	37,5
II.2	17	1	0,31	0,479	31,2	68,8
II.3	17	1	0,06	0,250	6,2	93,8
II.4	17	1	0,06	0,250	6,2	93,8
III.1	17	1	0,88	0,342	87,5	12,5
III.2	17	1	0,69	0,479	68,8	31,2
III.3	17	1	0,63	0,500	62,5	37,5
III.4	17	1	0,81	0,403	81,3	18,7
III.5	17	1	0,88	0,342	87,5	12,5
III.6	17	1	1,00	0,000	100,0	0,0

Tab. 26a: Einschätzung der Lehrkräfte zu den bereitgestellten Kenntnissen und Fähigkeiten hinsichtlich der Testaufgaben; fettgedruckte Prozentangaben: über die Hälfte der beteiligten Lehrkräfte geben an, die Kenntnisse und Fähigkeiten im Unterricht bereitgestellt zu haben.

Aufgabe	N	fehlende Angaben	M	SD	eher ja (%)	eher nein (%)
Art der Aufgabe im Unterricht von Bedeutung						
I.1	17	1	0,87	0,352	86,7	13,3
I.2	17	1	0,19	0,403	18,8	81,2
I.3a	17	1	0,38	0,500	37,5	62,5
I.3b	17	1	0,38	0,500	37,5	62,5
I.4	17	1	0,69	0,479	68,8	31,2
I.5a	17	1	0,94	0,250	93,7	6,3
I.5b	17	1	0,69	0,479	68,8	31,2
I.5c	17	1	0,69	0,479	68,8	31,2
II.1a	17	1	0,81	0,403	81,3	18,7
II.1b	17	1	0,81	0,403	81,3	18,7
II.2	17	1	0,50	0,516	50,0	50,0
II.3	17	1	0,31	0,479	68,8	31,2
II.4	17	1	0,25	0,447	25,0	75,0

Aufgabe	N	fehlende Angaben	M	SD	eher ja (%)	eher nein (%)
III.1	17	1	1,00	0,000	100,0	0,0
III.2	17	1	0,88	0,342	87,5	12,5
III.3	17	1	0,88	0,342	87,5	12,5
III.4	17	1	0,88	0,342	87,5	12,5
III.5	17	1	1,00	0,000	100,0	0,0
III.6	17	1	1,00	0,000	100,0	0,0

Tab. 26b: Einschätzung der Lehrkräfte zur Bedeutung der Aufgabenart für ihren Unterricht; fettgedruckte Prozentangaben: über die Hälfte der teilnehmenden Lehrkräfte geben an, dass die Art der Aufgabe für ihren Unterricht von Bedeutung ist.

In den Tabellen sind die Aufgabennummern und die entsprechenden Prozentzahlen hervorgehoben, wenn über die Hälfte der Lehrkräfte aussagen, dass sie die für die Bearbeitung der Aufgabe benötigten Kenntnisse und Fähigkeiten im Unterricht bereitgestellt haben (Tab. 26a) und dass diese Art der Aufgaben für ihren Unterricht generell von Bedeutung sind (Tab. 26b).

Es zeigt sich, dass alle Aufgaben aus dem Testteil „Schulchemie“ für die Lehrkräfte, die an der Untersuchung teilgenommen haben, von großer Bedeutung sind (im Durchschnitt 87,5 bzw. 100 %) und dass sie für diese sechs Aufgaben die Bereitstellung der erforderlichen Kenntnisse und Fähigkeiten mit durchschnittlich 62,5 – 100 % hoch einschätzen. Bei den anderen beiden Testteilen fällt auf, dass die Aufgaben, die sich dadurch auszeichnen, dass sie verstärkt das Alltagswissen der Lernenden aufgreifen, wie z. B. das Formulieren von Kriterien, Fragen und Internet-Suchbegriffen (zur Klassifizierung der Aufgaben siehe Tab. 13a und 13b, S. 68, 69), zu einer Art von Aufgaben zählen, die nur für wenige der teilnehmenden Lehrkräfte eine Bedeutung im Unterricht besitzen und für die entsprechend nur von wenigen Lehrkräften die nötigen Fertigkeiten bereitgestellt wurden. An späterer Stelle werden diese Aussagen der Lehrkräfte und zusätzlich die Einschätzungen zum Abschneiden ihrer Klasse bei exemplarisch betrachteten Aufgaben mit dem tatsächlichen Abschneiden der Klassen in Zusammenhang gebracht.

6.4.2 Zwischendiskussion 1

Zusammenfassend lässt sich festhalten, dass eine Passung der im Fragebogen erfassten Inhalte zum Unterricht nach Einschätzung der Lehrkräfte gegeben ist. Für zwölf der 19 Testitems geben über 50 % der Lehrkräfte an, die erforderlichen Fertigkeiten im Unterricht vermittelt zu haben, und zudem sind nach Auffassung von über der Hälfte der Lehrkräfte 14 der 19 Aufgaben generell von Bedeutung für ihren Chemieunterricht. Nach diesen Werten sind die Lernenden aus fachlicher Sicht in der Lage, die Aufgaben zu beantworten. Die nach Aussage der Lehrkräfte nicht bereitgestellten Kenntnisse und Fähigkeiten beziehen sich auf Aufgaben, die in der Klassi-

fizierung überwiegend dem Bereich des „Alltagswissens“ zugeordnet sind (vgl. Tab. 13a, S. 68). Dennoch ist es erstaunlich, dass trotz der hohen Beteiligung von Lehrkräften, die die Konzeption *Chemie im Kontext* kennen oder sogar in Schulsets mitarbeiten (zwölf der 17 teilnehmenden Lehrkräfte), bestimmte Aufgabenarten, die entsprechend den Zielen der Konzeption eher typisch sind, wenig Bedeutung im Unterricht haben (z. B. Formulieren von Fragen, Suchbegriffen und Kriterien, Bewerten der Aussagekraft einer Studie). Dass die Aufgaben, die dem dritten Fragebogenteil entstammen, durchweg in ihrer Bedeutung im Unterricht als wichtig beurteilt werden, ist weniger überraschend, denn die traditionellen Aufgaben bleiben offensichtlich auch in einem Unterricht nach *Chemie im Kontext* weiterhin zentral.

6.4.3 Ergebnisse der Auswertung des Schülerfragebogens

In der folgenden Ergebnisdarstellung werden zunächst die Erkenntnisse zum ersten Fragenkomplex zur statistischen Prüfung des Testinstruments (Fragestellung 1) vorgestellt. Im Anschluss daran werden die Leistungen der Schülerinnen und Schüler thematisiert (Fragestellung 2) und abschließend werden Ergebnisse der Einzelbetrachtungen ausgewählter Testaufgaben (Fragestellung 3) berichtet.

6.4.3.1 Statistische Prüfung des Testinstruments (Fragestellung 1)

(1) Fragestellung 1.1:

Reliabilitätsanalyse: Mit welchen Reliabilitätswerte erfasst das Untersuchungsinstrument die Leistungen der Schülerinnen und Schüler

Zur Beantwortung dieser Fragestellung wird zunächst berichtet, mit welcher Reliabilität der gesamte Test die Leistungen der Schülerinnen und Schüler ermittelt. Die Berechnung zeigt, dass bei einer Berücksichtigung aller 19 Testitems die interne Konsistenz des Tests bei Cronbachs $\alpha = 0,72$ liegt, wobei jedoch angemerkt werden muss, dass drei Items zu geringe Trennschärfen besitzen. Als nicht trennscharf in Bezug auf die durch die Testitems gemessene Dimension erweisen sich die Aufgaben I.1 („Gegenmittel Sodbrennen“), II.3 („Wirkungsweise von Zahnpflegekaugummi“) und II.4 („Studie bewerten“).

Bleiben diese drei Aufgaben in der Reliabilitätsanalyse unberücksichtigt, lässt sich ein Wert für die interne Konsistenz des Fragebogens von Cronbachs $\alpha = 0,74$ für die restlichen 16 Testitems ermitteln, dabei liegen die Trennschärfen von sechs Items $> 0,20$, die der restlichen zehn $> 0,30$. Die entsprechenden Kenndaten können den Tabellen 27 und 28 entnommen werden.

Testaufgabe	N	M	SD	Trennschärfe
I.1*	355	1,30	0,66	0,07
I.2*	337	1,28	0,86	0,37
I.3a*	351	1,27	0,71	0,20
I.3b*	353	1,56	0,73	0,26
I.4*	281	1,47	1,13	0,28
I.5a	287	0,68	0,47	0,51
I.5b*	196	0,54	0,84	0,38
I.5c*	297	1,05	0,51	0,43
II.1a	333	0,33	0,47	0,23
II.1b*	341	1,34	0,83	0,29
II.2	272	0,56	0,50	0,29
II.3*	351	0,92	0,84	0,12
II.4	355	0,49	0,50	0,06
III.1	302	0,44	0,50	0,34
III.2	332	0,72	0,45	0,39
III.3*	320	0,98	0,63	0,40
III.4	334	0,83	0,38	0,38
III.5*	251	0,96	0,92	0,36
III.6*	251	0,86	0,82	0,43

Tab. 27: Kenndaten zur internen Konsistenz des Gesamttests; **Cronbachs $\alpha = 0,72$** ; für die Bestimmung der Trennschärfen sind fehlende Antworten als falsch gewertet worden, daher $N = 359$; für die in der Tabelle angegebenen Kenndaten werden die Werte für die tatsächlich bearbeiteten Aufgaben (N) wiedergegeben, die mehrstufig ausgewerteten Aufgaben sind durch ein Sternchen gekennzeichnet.

Testaufgabe	N	M	SD	Trennschärfe
I.2*	337	1,28	0,85	0,37
I.3a*	351	1,27	0,71	0,21
I.3b*	353	1,56	0,73	0,28
I.4*	281	1,47	1,13	0,29
I.5a	287	0,68	0,47	0,52
I.5b*	196	0,54	0,84	0,40
I.5c*	297	1,05	0,51	0,44
II.1a	333	0,33	0,47	0,23
II.1b*	341	1,34	0,83	0,28
II.2	272	0,56	0,50	0,29
III.1	302	0,44	0,50	0,34
III.2	332	0,72	0,45	0,38
III.3*	320	0,98	0,63	0,38
III.4	334	0,83	0,38	0,39
III.5*	251	0,96	0,92	0,38
III.6*	251	0,86	0,82	0,43

Tab. 28: Kenndaten zur internen Konsistenz des Test, wenn nur die 16 Items berücksichtigt werden, deren Trennschärfe $> 0,20$ ist; **Cronbachs $\alpha = 0,74$** ; für die Bestimmung der Trennschärfen sind fehlende Antworten als falsch gewertet worden, daher $N = 359$; für die in der Tabelle angegebenen Kenndaten werden die Werte für die tatsächlich bearbeiteten Aufgaben (N) wiedergegeben, die mehrstufig ausgewerteten Aufgaben sind durch ein Sternchen gekennzeichnet.

(2) Fragestellung 1.2

Faktorenanalyse: Ist es anhand des Tests möglich, entsprechend der zu Grunde gelegten Intention der Erfassung verschiedener Aspekte des naturwissenschaftlichen Arbeitens verschiedene Skalen abzubilden?

Im Folgenden wird untersucht, ob und inwieweit es anhand der Testitems gelingt, verschiedene Komponenten des Leistungsspektrums zu erfassen. Dieses wird mithilfe des Verfahrens der Faktorenanalyse untersucht, bei der Faktorenanalyse wird als Extraktionsmethode die Hauptkomponentenanalyse, als Rotationsverfahren die Varimaxrotation gewählt.

Bei Betrachtung des Eigenwerteverlaufs (Screeplot: siehe Anhang I, Abb. 19, S. 210) wird deutlich, dass der größte Varianzanteil der Testitems durch drei bis fünf Faktoren erklärt werden kann. Für die Faktorenanalyse wird von drei Faktoren ausgegangen, in die Berechnung fließen die 16 Testitems ein, die in der Auswertung des Gesamttests aufgrund geeigneter Trennschärfen berücksichtigt werden. Die rotierte Komponentenmatrix ist im Anhang I (Tab. 51, S. 210) wiedergegeben.

Die Faktorenanalyse liefert für eine Mehrdimensionalität des Messinstruments weder in Hinblick auf die zu Grunde gelegten Kompetenzen noch auf die kognitive Aufgabenanforderung Anhaltspunkte. Es lassen sich zwar Faktoren ermittelt, aber zum einen besitzen mehrere Aufgaben starke Nebenladungen und zum anderen weisen die internen Konsistenzen dieser Skalen geringe Werte auf (vgl. Anhang I, Tab. 51, S. 210).

(3) Fragestellung 1.3

Schwierigkeitsanalyse: Durch welche Schwierigkeitsindices sind die Testaufgaben gekennzeichnet?

Für die Aufgaben des Fragebogens der Hauptuntersuchung werden, wie bereits für die Testaufgaben der Pilotstudie, die Schwierigkeiten bestimmt, indem nach den im Zusammenhang mit der Fragestellung 2.4 der Pilotstudie angegebenen Formeln die Schwierigkeitsindices für die Items bestimmt werden (vgl. Kap. 5.4.3 dieser Arbeit).

Die durchschnittliche Itemschwierigkeit des Tests der Hauptuntersuchung beträgt $p_{\text{ges}} = 0,55$ und ist damit insgesamt etwas leichter als die des Tests der Pilotstudie ($p_{\text{ges}} = 0,48$). Die Schwierigkeiten der 19 Testaufgaben können Tabelle 29 entnommen werden. In der Tabelle gibt die Zahl hinter der jeweiligen Aufgabenanforderung die Anzahl der für das jeweilige Item zugelassenen teilrichtigen Antworten an.

Insgesamt haben sich fünf Testaufgaben als einfach, elf Items als mittelschwer und drei Items als schwierig erwiesen. Es ist folglich gelungen, den Anteil der schwierigen Aufgaben zu senken und den der mittelschweren Aufgaben zu erhöhen.

Test-aufgabe	Aufgabenanforderung, (Antwortstufen)	gültige Antworten	fehlende Antworten	M	SD	p
<i>Aufgaben mit niedrigem Schwierigkeitsgrad</i>						
I.3	b) Formulieren von Suchbegriffen für eine Internetrecherche (2)	353	6	1,56	0,73	0,78
I.5	a) Aufstellen einer Reaktionsgleichung	287	72	0,68	0,47	0,68
II.1	b) Interpretieren der Grafik (2)	341	18	1,34	0,83	0,67
III.2	Identifizieren der pH-Kurve, die die beschriebene Neutralisationstitation korrekt wiedergibt	332	27	0,72	0,45	0,72
III.4	Auswählen eines geeigneten Indikators aus gegebener Liste	334	25	0,83	0,38	0,83
<i>Aufgaben mit mittlerem Schwierigkeitsgrad</i>						
I.1.	Auswählen einer Substanz als Gegenmittel bei Sodbrennen (2)	355	4	1,30	0,66	0,65
I.2.	Nennen wichtiger Kriterien bei der Medikamententwicklung (2)	337	22	1,28	0,85	0,64
I.3	a) Formulieren von adressatengerechten Fragen (2)	351	8	1,27	0,71	0,63
I.4	Erstellen einer Versuchszeichnung (3)	281	78	1,47	1,13	0,49
I.5	c) weitere Informationen einholen (2)	297	62	1,05	0,51	0,52
II.2	Identifizieren von fehlerhaften Informationen in der grafischen Auftragung	272	87	0,56	0,50	0,55
II.3	Beurteilen der Wirkmechanismen eines Zahnpflegekaugummi (2)	351	8	0,92	0,84	0,46
II.4	Beurteilen der Aussagekraft einer vorgestellten Studie	355	4	0,49	0,50	0,49
III.1	Identifizieren einer Stoffaufzählung, die ausschließlich Stoffe nennt, die in Wasser gelöst basisch reagieren	302	57	0,44	0,50	0,44
III.3.	Treffen einer begründeten Entscheidung für eine Vorgehensweise bei verschütteter Säure (2)	320	39	0,98	0,63	0,49
III.6	Erklären des Phänomens der elektrischen Leitfähigkeit (2)	251	108	0,86	0,82	0,43
<i>Aufgaben mit hohem Schwierigkeitsgrad</i>						
I.5	b) Durchführen einer stöchiometrischen Berechnung (2)	196	163	0,54	0,84	0,27
II.1	a) Beschreiben eines Kurvenverlaufs	333	26	0,33	0,47	0,33
III.5.	Nennen von Eigenschaften saurer Lösungen mit geeigneten Demoversuchen (3)	251	108	0,96	0,92	0,32

Tab. 29: Schwierigkeitsindices der einzelnen Testaufgaben. In der Tabelle gibt die Zahl hinter der jeweiligen Aufgabenanforderung die Anzahl der für das jeweilige Item zugelassenen Antwortalternativen an, darüber hinaus angegeben Mittelwerte und Standardabweichungen.

Im Vergleich zu den Itemschwierigkeiten des Fragebogens der Pilotuntersuchung lässt sich feststellen, dass es gelungen ist, durch die Veränderungen im Erwartungshorizont den Schwierigkeitsgrad einzelner Aufgaben zu heben (I.3b „Internet-Suchbegriffe“) bzw. zu senken (III.6 „Leitfähigkeit“), durch die Einführungen einer (weiteren) teilrichtigen Lösung (I.4 „Versuch zum Gas auffangen“, II.3 „Wirkungsweise von Zahnpflegekaugummi“) oder durch Konkretisierung (II.1 und II.2 „Beschreibung und Interpretation der Kurve“, III.6 „Leitfähigkeit“) bzw. Vereinfachung (I.5b „Berechnung“) der Aufgabenstellung den Schwierigkeitsgrad zu senken. Bei der Aufgaben III.5 sind zwei Veränderungen vorgenommen worden, zum einen ist das Vorgehen in der Aufgabenstellung konkretisiert worden, zum anderen ist die Dreistufigkeit des Antwortformat zu einer vierstufigen Skala hin verändert worden. Dennoch fällt es den Probanden der Hauptstudie schwerer, diese Aufgabe erfolgreich zu bearbeiten.

Für die übrigen Items kann berichtet werden, dass sich die Aufgabenschwierigkeiten nur in wenigen Fällen deutlich verändert haben: Das Formulieren von adressatengerechten Fragen fällt der Stichprobe der Hauptuntersuchung immer noch leicht, aber dennoch schwerer als den Probanden der Pilotstudie ($p = 0,63$ in der Hauptstudie gegenüber $p = 0,79$ in der Pilotstudie). Demgegenüber fällt den Schülerinnen und Schülern der Hauptstudie das Formulieren der Reaktionsgleichung bedeutend leichter ($p = 0,68$) als den Teilnehmern der Pilotstudie ($p = 0,43$). Die Aufgaben, die den Untersuchungsteilnehmern der Hauptstudie die größten Schwierigkeiten bereiten, sind durch folgende Anforderungen gekennzeichnet: zum einen müssen die Probanden eine einfache stöchiometrische Berechnung anstellen ($p = 0,27$), zum anderen müssen sie den Verlauf einer Grafik beschreiben ($p = 0,33$) und Eigenschaften saurer Lösungen mit geeigneten Demonstrationsversuchen in einer Tabelle notieren ($p = 0,32$).

Bei mehrstufigen Items wird zusätzlich zur Itemschwierigkeit die Lösungshäufigkeit der einzelnen Antwortstufen betrachtet (dargestellt in Tab. 30). Ziel ist es, durch die Erfassung von teilrichtigen Antworten, detailliertere Aussagen über Schwierigkeiten der Schülerinnen und Schüler beim Bearbeiten einzelner Aufgaben treffen zu können.

Testaufgaben	Antwortalternativen	Anzahl Antworten	gültige Prozent
	a) vollst. u. korrekt, b) teilrichtig 1 (ggf. teilrichtig 2), c) falsch	Σ gesamt	
I.1.	a) Alternative „basisches Salz“ angekreuzt	144	40,6
	b) Alternative „basisches Salz“ und „Lauge“ oder nur Lauge“ angekreuzt	172	48,5
	c) falsche Alternative	39	11,0
		Σ 355	
I.2	a) nennt drei sinnvolle Kriterien	176	52,2
	b) nennt zwei sinnvolle Kriterien	81	24,0
	c) nennt weniger als zwei sinnvolle Kriterien	80	23,7
		Σ 337	
I.3	a) formuliert drei adressatengerechte Fragen	147	41,9
	b) formuliert zwei adressatengerechte Fragen	151	43,0
	c) formuliert nur eine adressatengerechte Frage	53	15,1
		Σ 351	
	a) formuliert drei geeignete Suchbegriffe	247	70,0
	b) formuliert zwei geeignete Suchbegriffe	55	15,6
	c) formuliert nur einen geeigneten Suchbegriff	51	14,4
		Σ 353	
I.4	a) Zeichnung gibt einen sinnvollen Aufbau vollständig wieder	68	24,2
	b) Zeichnung gibt einen sinnvollen Aufbau mit kleineren Schwächen wieder	73	26,0
	c) insgesamt richtige Zeichnung, aber geschlossene Apparatur	64	22,8
	d) zu unvollständig, falsch	76	27,0
		Σ 281	
I.5b	a) vollständige und richtige Berechnung	45	23,0
	b) richtige Berechnung der molaren Massen	16	8,2
	c) falsche Rechnung	135	68,9
		Σ 196	
I.5c	a) erfragt u. a. den Säureüberschuss im Magen	45	15,2
	b) erfragt andere Informationen	221	74,4
	c) erfragt unsinnige Informationen	31	10,4
		Σ 297	
II.1b	a) interpretiert anhand der Kurve	195	57,2
	b) greift auf Informationen aus dem Text zurück	68	19,9
	c) liefert eine falsche Interpretation	78	22,9
		Σ 341	
II.3	a) beurteilt alle vier Aussagen korrekt	109	31,1
	b) beurteilt drei Aussagen korrekt	105	29,9
	c) beurteilt zwei oder weniger Aussagen korrekt	137	39,0
		Σ 351	
III.3	a) entscheidet sich mit richtiger Begründung	59	18,4
	b) entscheidet sich für Natronlauge, weil es neutralisiert	195	60,9
	c) begründet Entscheidung nicht oder falsch	66	20,6
		Σ 320	
III.5	a) nennt zwei Eigenschaften mit je einem Demoversuch	20	8,0
	b) füllt drei Felder der Tabelle richtig aus	42	16,7
	c) füllt zwei Felder der Tabelle richtig aus	97	38,6
	d) nennt Falsches und damit insgesamt weniger als unter c)	92	36,7
		Σ 251	
III.6	a) erklärt Leitfähigkeit, geht dabei auf saure Lösungen und dest. Wasser ein	69	27,5
	b) erklärt Leitfähigkeit generell, geht dabei nicht oder nur z. T. auf saure Lösung/dest. Wasser ein	78	31,1
	c) falsche Erklärung	104	41,4
		Σ 251	

Tab. 30: Lösungshäufigkeiten der einzelnen Antwortstufen bei mehrstufigen Items.

Tabelle 30 bestätigt, dass durch das Zulassen einer (weiteren) teilrichtigen Antwort die Schwierigkeit gesenkt wird, da auf diese Weise eine weitere Gruppe von Schülerinnen und Schülern in der Auswertung erfasst wird. Für Aufgabe III.5 muss jedoch festgehalten werden, dass es in diesem Fall nicht zu der gewünschten Vereinfachung geführt, denn nur ein geringer Prozentsatz von 8 % der Schülerinnen und Schüler ist in der Lage, alle vier Felder der Tabelle korrekt zu füllen. Die gezielte Erfassung von Versuchszeichnungen, die eine geschlossene Gas auffangapparatur widerspiegeln (Aufgabe I.4), zeigt, dass etwa ein Fünftel der teilnehmenden Schülerinnen und Schüler (22,8 %) eine insgesamt sinnvolle, aber geschlossene Apparatur entwerfen. Im Rahmen der Fragestellung 3 wird auf diese Tabelle zur vertieften Betrachtung ausgewählter Aufgaben zurückgegriffen.

6.4.3.2 Zwischendiskussion 2

Reliabilität und Dimensionalität der Testaufgaben

Zusammenfassend ist festzustellen, dass sich der Fragebogen als Instrument zur Erfassung von Leistungen der Schülerinnen und Schüler insgesamt eignet. Die interne Konsistenz des Fragebogens hat sich im Vergleich zum Fragebogen der Pilotstudie (Cronbachs $\alpha = 0,77$) zwar insgesamt etwas verringert (Cronbachs $\alpha = 0,74$), aber bedeutend weniger Aufgaben besitzen zu geringe Trennschärfen, um sie in der Gesamtbetrachtung berücksichtigen zu können. Während sich die Skala des Fragebogens der Pilotuntersuchung aus zehn der 17 Items zusammensetzt, erfasst der Fragebogen der Hauptuntersuchung eine Dimension, die aus 16 Aufgaben besteht. Die durchgeführte Faktorenanalyse liefert keine Hinweise auf eine Mehrdimensionalität des Untersuchungsinstruments, vielmehr muss davon ausgegangen werden, dass die einzelnen Facetten des naturwissenschaftlichen Arbeitens, wie sie in Kapitel 4 detailliert beschrieben sind, anhand dieser 16 Aufgaben nicht getrennt voneinander erfasst werden können. Sie bestätigt damit das Ergebnis der Pilotstudie hinsichtlich der Dimensionalität des Testinstruments. Die Aufgaben des Fragebogens messen insgesamt Kenntnisse und Fähigkeiten der Probanden auf dem Themengebiet „Säuren, Basen und Neutralisation“ mit einerseits inhaltlicher Ausrichtung auf Alltagsthemen und andererseits inhaltlicher Orientierung an klassischen Themenaspekten aus dem Chemieunterricht.

Die anhand des Fragebogens erfasste Dimension erstreckt sich über verschiedene Aspekte des Säure-Base-Konzepts über einfaches Erkennen von geeigneten Indikatoren für einen bestimmten pH-Bereich und das Aufstellen einer Reaktionsgleichungen einer Neutralisationsreaktion bis hin zur Erklärung des Phänomens der elektrischen Leitfähigkeit anhand von Modellvorstellungen. Darüber hinaus wird der gemessene Bereich durch verschiedene Kompetenzen beschrieben, die

sich aus den beschriebenen Schritten bei der Erarbeitung eines Kontextes ergeben. Es werden Kompetenzen erfasst, die eher dem Bereich der Kontexterschließung und der Planungsphase zugeordnet werden können, wie das Formulieren von Kriterien (I.2) und (weiterführenden) Fragen (I.3a, I.5c), aber auch Fähigkeiten, die der Bereich der Anwendung und Einordnung von Kenntnissen in den Kontext (z. B. II.2) bzw. in den fachchemischen Zusammenhang (z. B. III.2) erfordert. Darüber hinaus werden in den Aufgaben verschiedene Anforderungen an die Probanden gestellt, die unterschiedliche Anforderungen im gesamten Bereich des naturwissenschaftlichen Arbeitens darstellen (siehe Tab. 13a und 13b, S. 68, 69).

Hiermit bestätigt sich die Schwierigkeit des getrennten Erfassens einzelner Kompetenzen, da sie sich offenbar gegenseitig bedingen und zudem mit den erforderlichen Wissensbeständen in engem Zusammenhang stehen.

Itemschwierigkeiten

Mithilfe der berechneten Schwierigkeitsindices der Testaufgaben lässt sich feststellen, welche Schritte und Anforderungen des naturwissenschaftlichen Arbeitens im Themenbereich des Säure-Base-Konzepts den Schülerinnen und Schülern eher leicht fallen und welche ihnen eher Schwierigkeiten bereiten. Wenig Probleme treten bei den Aufgaben I.3 („Internet-Suchbegriffe“), III.2 („pH-Kurven“) und III.4 („Indikator“) auf, schwer fallen den Schülerinnen und Schülern die Aufgaben I.5b („Berechnung“), II.1a („Kurve beschreiben“) und III.5 („Demonstrationsversuche“) (Tab. 29). Auffallend ist, dass sich die drei als leicht klassifizierten Aufgaben von den als schwierig eingestuften Aufgaben hinsichtlich des geforderten Antwortformats deutlich unterscheiden. Während die Probanden für die drei schwierigen Aufgaben die Antwort frei formulieren müssen, handelt es sich bei den Antwortformaten der einfachen Aufgaben in zwei Fällen um ein Multiple-Choice-Format und im dritten Fall müssen drei Begriffe genannt werden. Festgehalten werden kann zudem, dass es den Teilnehmern der Studie nicht besonders schwer fällt, im fachlichen Kontext die Aussage von pH-Wert-Kurven richtig zu erfassen bzw. zuzuordnen, dass es ihnen aber sehr wohl Schwierigkeiten bereitet, eine vorgegebene, inhaltlich in den Kontextaspekt „Kaugummi“ eingebetteten Kurvenverlauf mit eigenen Worten zu beschreiben. Die schwierigste Aufgabe des Tests erfordert von den Probanden eine einfache stöchiometrische Berechnung, an der sich insgesamt nur 196 Schülerinnen und Schüler überhaupt versuchen, von denen 61 (entspricht 31 %) eine teilrichtige oder richtige Lösung produzieren, obwohl 62,5 % der unterrichtenden Lehrkräfte angeben, die Fähigkeiten zum Lösen dieser Aufgaben im Unterricht bereitgestellt zu haben.

6.4.3.3 Leistungen der Schülerinnen und Schüler (Fragestellung 2)

(1) Fragestellung 2.1

Gesamtleistungen: Welche Erkenntnisse liefert der Fragebogen in Bezug auf die unterschiedlichen Leistungen der Schülerinnen und Schüler?

Betrachtet wird zunächst der Gesamttest (bestehend aus 16 Testaufgaben) in Hinblick auf das Abschneiden der Probanden insgesamt. Um wiederum die Gewichtung der einzelnen Items zu berücksichtigen, wird für jede Aufgabe als höchste Punktzahl 1 festgelegt, teilrichtige Antworten erhalten entsprechend geringere Werte. Als weitere Vereinfachung wird eine Zusammenfassung der erreichten Punktzahlen vorgenommen, so dass in der Grafik die Häufigkeiten für ganze Punktzahlen aufgezeigt wird (erreichte Punktzahl liegt zwischen $0 \leq 1$ erfasst in Säule 1, erreichte Punktzahl zwischen $4 \leq 5$ zusammengefasst in Säule 5 usw.), so dass sich für die Darstellung insgesamt 16 Säulen ergeben.

Nach dieser Modifizierung der Daten ergibt sich für die betrachtete Gesamtmenge der Aufgaben eine maximal erreichbare Punktzahl von 16, die Verteilung aller Probanden nach ihren erreichten Punktzahlen stellt sich grafisch folgendermaßen dar (Abb. 8):

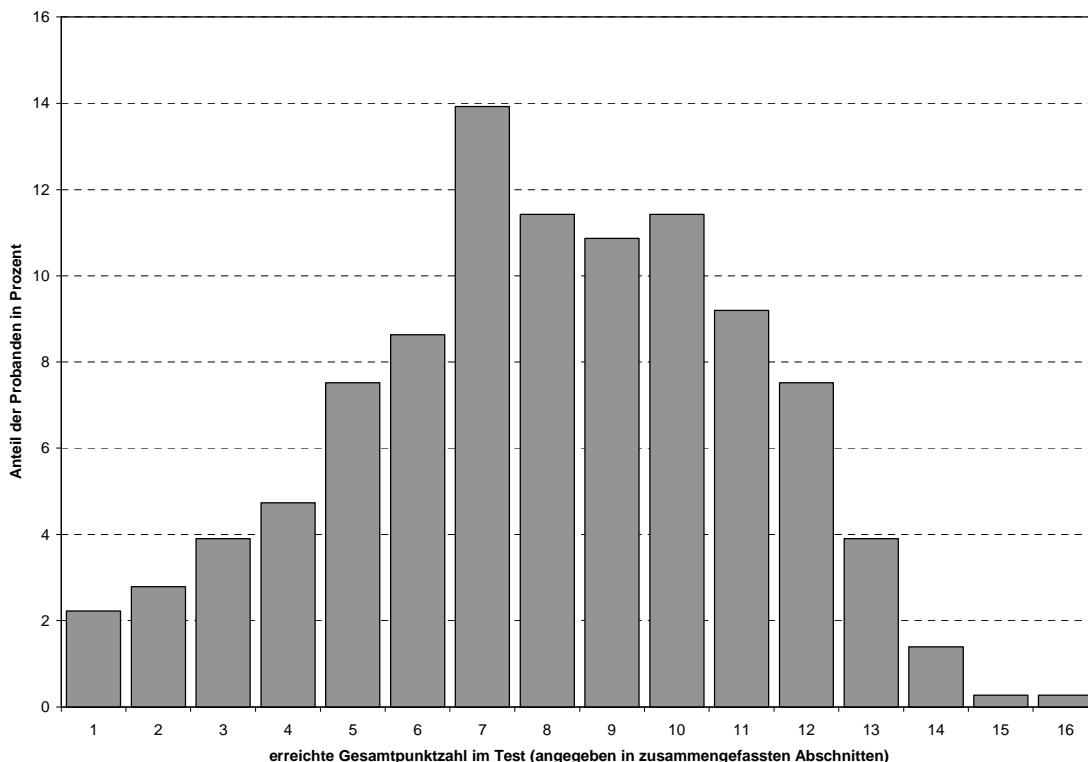


Abb. 8: Verteilung aller Probanden hinsichtlich ihrer erreichten Punktzahl im Gesamttest; $N = 359$, $M = 7,63$ bei einer Standardabweichung von $3,046$; erreichte Punktzahlen sind in Abschnitten zusammengefasst.

Die maximal erreichbare Gesamtpunktzahl von 16 wird von einem Teilnehmer der Untersuchung erreicht, keine Punkte bei den betrachteten 16 Testaufgaben erlangen zwei Probanden (N = 359, M = 7,63 bei einer Standardabweichung von 3,046)

In der folgenden Abbildung 9 wird das Abschneiden der 17 teilnehmenden Klassen hinsichtlich der 16 betrachteten Items dargestellt, die entsprechenden Kenndaten zu den einzelnen Klassen sind dem Anhang I (Tab. 52, S. 211) zu entnehmen.

Die Darstellung zeigt, dass es im Unterschied zur Pilotstudie deutliche Unterschiede in den Klassenleistungen gibt, zum einen im Gesamtabschneiden, zum anderen finden sich einzelne Klassen, deren Leistungen mit einer Standardabweichung von 1,16 bis 1,27 (Klasse 10, 15 und 17) relativ gering streuen und damit als Klasse in ihren Leistungen vergleichsweise konsistent sind, während in anderen Klassen die Streubreiten mit einer Standardabweichung von 3,32 bzw. 3,35 (Klasse 2 und 5) verhältnismäßig hoch liegen; diese Klassen zeigen folglich ein inhomogenes Leistungsbild.

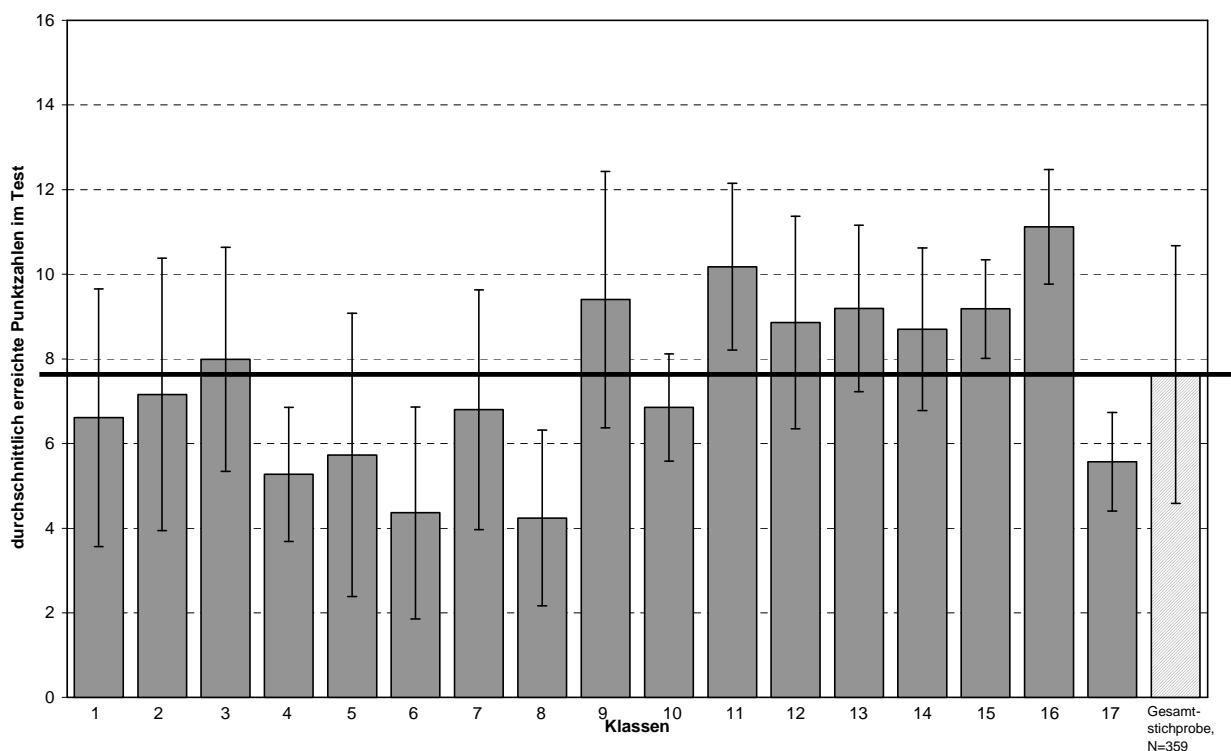


Abb. 9: Abschneiden der teilnehmenden Klassen hinsichtlich der 16 Items.

Darüber hinaus ist festzustellen, dass die teilnehmenden Klassen, die alle aus einem Bundesland stammen (Klasse 9 – 16) mit einer Ausnahme über dem Durchschnitt der Gesamtstichprobe liegen, während die Klassen aus den übrigen Bundesländern wiederum mit einer Ausnahme unter dem Gesamtdurchschnitt zu finden sind.

Eine einfaktorielle Varianzanalyse ergibt, dass es Unterschiede im Abschneiden der Klassen gibt, $F(16,342) = 17,82$, $p \leq 0,001$. Die Bestimmung von Scheffé-Kontrasten zeigt z. B., dass Klasse 16 signifikant besser abschneidet als die Klassen aus den anderen Bundesländern (Klasse 1, 2, 4, 5, 6, 7, 8 und 17) mit Ausnahme von Klasse 3 (bei einem betrachteten Signifikanzniveau $< 0,05$) und ebenfalls besser als Klasse 10, die aus dem gleichen Bundesland stammt. Der Mittelwert der erreichten Punktzahlen von Klasse 16 liegt mit 11,12 bei einer Standardabweichung von 1,36 deutlich über dem Gesamtdurchschnitt von 7,63 ($SD = 3,046$).

Besonders gravierende Unterschiede finden sich zwischen den im Vergleich leistungsstarken Klassen 9, 11 und 16 und den vergleichsweise leistungsschwachen Klassen 4, 6 und 8, wobei in allen Fällen die Klassen 9, 11 und 16 signifikant besser sind als die Klassen 4, 6 und 8 (Scheffé-Kontraste, Signifikanzniveau $< 0,05$). Die unter Fragestellung 2 vorgenommene Betrachtung des klassenspezifischen Antwortverhaltens sollen weiteren Aufschluss über besondere Stärken bzw. Schwächen einzelner Klassen in Hinblick auf besondere Aufgabenanforderungen geben. Diese Erkenntnisse werden mit den Einschätzungen, die von den jeweils unterrichtenden Lehrkräften vorgenommen werden, in Zusammenhang gebracht.

Gesonderte Betrachtung der Teilnehmer eines Bundeslandes

Die Stichprobe des Bundeslandes, aus dem besonders zahlreiche Teilnehmer der Untersuchung stammen, wird im Folgenden genauer untersucht, da die Klassen in ihren Erfahrungen mit der Unterrichtskonzeption *Chemie im Kontext* unterschieden werden können. Darüber hinaus ist in allen Klassen das Thema Säuren und Basen nicht nach *Chemie im Kontext* unterrichtet worden, sondern es wurden nur, wie in Tabelle 22 (Kap. 6.3.1, S. 109) dargestellt, beispielhaft Bezüge zur Alltagswelt der Lernenden hergestellt. Dennoch unterscheiden sich die Klassen darin, dass ein Teil der Probanden (Klasse 11, 12, 13, 14 und 15) in den vergangenen zwei Jahren mit drei bis fünf Kontexteinheiten unterrichtet worden ist, während den anderen Klassen (Klasse 9, 10 und 16) die Konzeption nicht bekannt ist. Es muss bei Betrachtung der Ergebnisse jedoch berücksichtigt werden, dass das entwickelte Instrument in der vorliegenden Arbeit nicht in erster Linie zur Evaluation der Unterrichtskonzeption *Chemie im Kontext* eingesetzt worden ist, sondern es vielmehr um eine kompetenzorientierte Lernstandstandserhebung (z. B. im Sinne von Vergleichsarbeiten) geht. Aus diesem Grund wird auf eine systematisierende Betrachtung des tatsächlichen Unterrichts verzichtet.

Es wird untersucht, ob sich Unterschiede im Abschneiden der beiden Gruppen in Hinblick auf den Lernstand zu Beginn der 11. Klasse feststellen lassen. Das Abschneiden der beiden Gruppen

im Gesamttest zeigt die folgende Abbildung 10, die entsprechenden Kennwerte können Tabelle 31 entnommen werden.

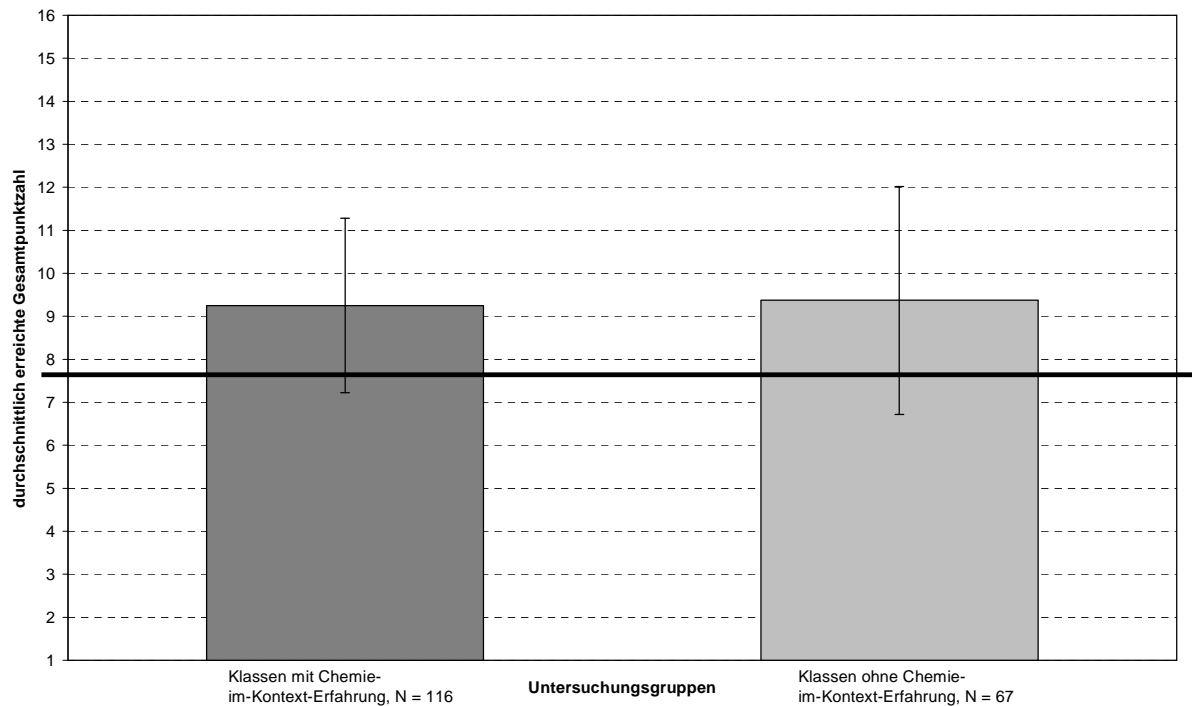


Abb. 10: Vergleich der beiden Gruppen mit und ohne *Chemie-im-Kontext*-Erfahrung hinsichtlich ihrer Gesamtleistung im Test (dabei jeweils betrachtet: 16 Items).

	N	M	SD
Probanden mit <i>Chemie-im-Kontext</i> -Erfahrung	116	9,25	2,029
Probanden ohne <i>Chemie-im-Kontext</i> -Erfahrung	67	9,34	2,651

Tab. 31: Kennwerte der beiden untersuchten Gruppen eines Bundeslandes.

Es zeigt sich kein signifikanter Unterschied in dem Abschneiden der beiden untersuchten Gruppen ($t_{(110,94)} = -0,33$, n.s.). Die Wissensbestände und Kompetenzen, die der Fragebogen zum Themengebiet „Säuren, Basen und Neutralisation“ erfasst, sind in beiden Untersuchungsgruppen gleichermaßen vorhanden. Es gibt demnach zwar Unterschiede zwischen den Klassen, aber nicht zwischen den anhand der vorliegenden *Chemie-im-Kontext*-Erfahrung gebildeten Gruppen.

Vergleich beider Gruppen insgesamt

Abschließend wird das Abschneiden der Probanden mit *Chemie-im-Kontext*-Erfahrung (N = 226) mit den Leistungen der Probanden ohne Erfahrung mit der Konzeption (N = 133) insgesamt über alle Bundesländer hinweg verglichen. Dabei müssen jedoch die großen Unterschiede innerhalb der Klassen u. a. hinsichtlich Stundenzahl der Unterrichtseinheit und Erhebungszeitpunkt bei der Diskussion berücksichtigt werden. Die bereits näher untersuchten Klassen, die aus einem Bundesland stammen und besonders gut im Test abschneiden, beeinflussen das Ergebnis insgesamt nicht unterschiedlich stark, da sie für beide Untersuchungsgruppen etwa die Hälfte der Probanden stellen.

Die entsprechenden Kennwerte können Tabelle 32 entnommen werden, das Abschneiden beider Gruppen stellt sich, wie in Abbildung 11 gezeigt, dar:

	N	M	SD
Probanden mit <i>Chemie-im-Kontext</i> -Erfahrung	226	7,85	2,820
Probanden ohne <i>Chemie-im-Kontext</i> -Erfahrung	133	7,24	3,372

Tab. 32: Kennwerte der beiden untersuchten Gruppen der Gesamtstichprobe.

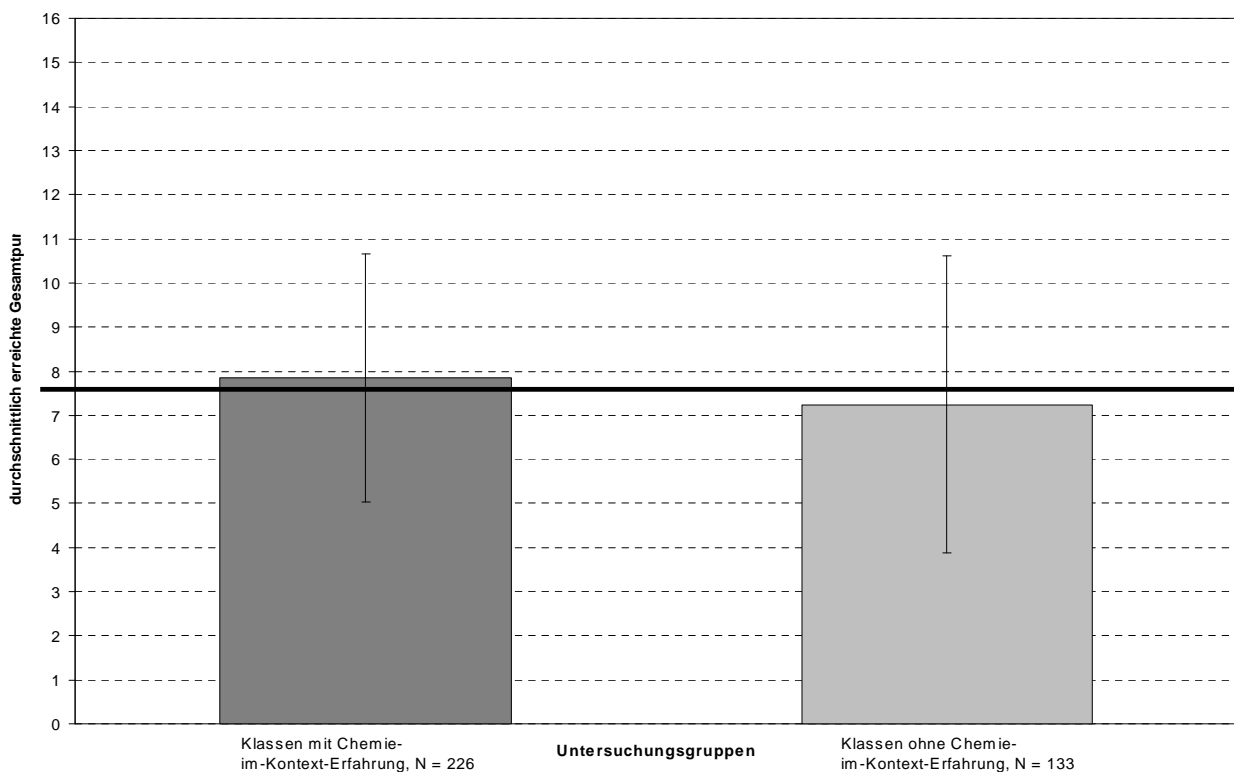


Abb. 11: Vergleich beider Gruppen (bundesländerübergreifend) mit und ohne *Chemie-im-Kontext*-Erfahrung hinsichtlich ihrer Gesamtleistung im Test (dabei jeweils betrachtet 16 Items).

Auch bei Betrachtung der gesamten Stichprobe lässt sich kein signifikanter Unterschied zwischen den Testergebnissen der Probanden, die nach *Chemie-im-Kontext*-Einheiten unterrichtet worden sind, und den Klassen, die keine Erfahrung mit der Konzeption haben, feststellen ($t_{(239,22)} = 1,76$, n.s.).

Ergebnisse aus dem Lehrerfragebogen

Tabelle 33 gibt Auskunft darüber, welche Kenntnisse und Fähigkeiten nach Einschätzung der unterrichtenden Lehrkräfte Thema des Unterrichts gewesen sind; in der Tabelle ist mit einem Kreuz gekennzeichnet, wenn die Aufgabenanforderungen im Unterricht thematisiert wurden.

Aufgabe	Klassen															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
I.1	x	x	x		x	x	x		x		x	x	x	x	x	x
I.2																
I.3a	x	x				x	x							x	x	
I.3b	x	x				x	x							x	x	
I.4	x	x			x	x	x		x		x	x	x	x	x	x
I.5a	x	x			x		x		x	x	x	x	x	k. A.	x	x
I.5b	x	x			x				x	x	x	x		x	x	x
I.5c		x			x	x					x			x	x	x
II.1a	x	x			x	x	x		x		x	x		x	x	
II.1b	x	x			x	x	x		x		x	x		x	x	
II.2	x				x							x		x	x	
II.3		x														
II.4																x
III.1	x	x	x	x	x	x		x	x	x	x	x	x		x	x
III.2		x	x	x	x	x		x	x	x				x	x	x
III.3		x	x		x	x	x				x		x	x	x	x
III.4	x	x		x	x		x	x	x	x	x	x	x		x	x
III.5	x	x	x		x	x	x		x	x	x	x	x	x	x	x
III.6	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Tab. 33: Lehrerfragebogen – Einschätzung der Lehrkräfte zur Bereitstellung der in den Aufgaben geforderten Kenntnissen und Fähigkeiten; Lehrkraft der Klasse 17 hat die Einschätzungen anhand des Fragebogens nicht vorgenommen; die drei in der Auswertung nicht berücksichtigten Aufgaben sind grau unterlegt.

Bei der Betrachtung des Testergebnisses einer Klasse und den nach Aussage der Lehrkraft thematisierten Kenntnissen und Fähigkeiten lässt sich kein Zusammenhang feststellen: Es gibt sowohl Klassen, in denen nach Angabe der Lehrkraft für mindestens zehn der insgesamt 16 be-

trachteten Aufgaben die benötigten Kenntnisse und Fähigkeiten bereitgestellt wurden, die aber im Gesamtabschneiden dennoch unter dem Durchschnitt liegen (Klassen 1, 2, 5, 6 und 7), als auch Klassen, in denen laut Lehrerfragebogen weniger als zehn Aufgabenanforderungen thematisiert wurden, die aber im Abschneiden über dem Durchschnitt liegen (Klassen 3 und 13). Die zu erwartenden Fälle, Aufgabeninhalte nicht im Unterricht berücksichtigt in Kombination mit einem eher schlechten Abschneiden und gutes Testergebnis bei unterrichtlicher Behandlung der Aufgabenanforderungen, treten ebenfalls auf (1. Fall: Klassen 4, 8 und 10; 2. Fall: Klassen 9, 11, 12, 14, 15 und 16).

(2) Fragestellung 2.2

Clusterzentrenanalyse: Durch welche Merkmale lassen sich die Aufgabengruppen beschreiben, die bei verschiedenen Probandengruppen ein unterschiedliches Antwortverhalten hervorrufen?

Die durchgeführte Faktorenanalyse zeigt, dass sich mit dem entwickelten Testinstrument einzelne Kompetenzen aus dem Bereich des naturwissenschaftlichen Denkens und Arbeiten nicht isoliert erfassen lassen, daher wird ermittelt, ob sich bei den Lernenden bestimmte Antwortmuster hinsichtlich der Aufgabenbearbeitung finden lassen.

Um zu untersuchen, ob sich die Probanden aufgrund ihres Antwortverhaltens in Hinblick auf eine Auswahl bestimmter Testaufgaben zu Gruppen zusammenführen lassen, wird eine Clusterzentrenanalyse durchgeführt. Mit diesem Verfahren werden „Fälle“ (hier die teilnehmenden Schülerinnen und Schüler) anhand ausgewählter Variablen in Gruppen (Cluster) eingeteilt, so dass die Variablen innerhalb einer Gruppe ähnliche Werte aufweisen, während sich die Variablenwerte verschiedenen Gruppen größtmöglich unterscheiden.

Für das Verfahren der Clusteranalyse werden aus den 19 Aufgaben des Gesamttests zehn ausgewählt, anhand derer zunächst mittels einer Faktorenanalyse vier Faktoren extrahiert werden.

Tabelle 34 zeigt, welche Testitems in die Clusteranalyse einfließen und welche unberücksichtigt bleiben; dabei wird die Entscheidung bei Nichtberücksichtigung jeweils kurz begründet, insgesamt wird versucht, möglichst aussagekräftige Aufgaben auszuwählen.

Tes- tauf- gabe	Aufgabenanforderung	fließt ein	unberück- sichtigt	Begründung
I.1.	Auswählen einer Substanz als Gegenmittel bei Sodbrennen		X	Formulierung der Aufgaben erweist sich als problematisch (Unklarheit über die Anzahl anzukreuzender Distraktoren).
I.2.	Nennen wichtiger Kriterien bei der Medikamententwicklung		X	Anforderung an das Verfassen von Texten ist im Vergleich niedrig, da nur einzelne Begriffe genannt werden müssen und wird daher zugunsten der Aufgaben I.3 nicht berücksichtigt.
I.3	a) Formulieren von adressatengerechten Fragen	X		
	b) Formulieren von Suchbegriffen für eine Internetrecherche		X	Aufgabe besitzt zu niedrigen Schwierigkeitsgrad.
I.4	Erstellen einer Versuchszeichnung	X		
I.5	a) Aufstellen einer Reaktionsgleichung	X		
	b) Durchführen einer stöchiometrischen Berechnung		X	Aufgabe besitzt zu hohen Schwierigkeitsgrad.
	c) weitere Informationen einholen		X	Nur ein sehr geringer Prozentsatz der Probanden antwortet entsprechend der Musterlösung für eine korrekte Lösung.
II.1	a) Beschreiben eines Kurvenverlaufs	X		
	b) Interpretieren der Grafik	X		
II.2	Identifizieren von fehlerhaften Informationen in der grafischen Auftragsung	X		
II.3	Beurteilen der Wirkmechanismen eines Zahnpflegekaugummi	X		
II.4	Beurteilen der Aussagekraft einer vorgestellten Studie		X	Es wird versucht, aus allen drei Testteilen eine gleiche Anzahl von Aufgaben einzubeziehen, zudem verleitet ein Distraktor eindeutig zu einer falschen Antwort.
III.1	Identifizieren einer Stoffaufzählung, die ausschließlich Stoffe nennt, die in Wasser gelöst basisch reagieren	X		
III.2.	Identifizieren der pH-Kurve, die die beschriebene Neutralisationstitation korrekt wiedergibt	X		
III.3.	Treffen einer begründeten Entscheidung für eine Vorgehensweise bei verschütteter Säure		X	Nur ein sehr geringer Prozentsatz der Probanden antwortet entsprechend der Musterlösung für eine korrekte Lösung.
III.4	Auswählen eines geeigneten Indikators aus gegebener Liste		X	Aufgabe besitzt zu niedrigen Schwierigkeitsgrad.
III.5.	Nennen von Eigenschaften saurer Lösungen mit geeigneten Demoversuchen		X	Aufgabenformat (tabellarische Form) erweist sich als problematisch.
III.6	Erklären des Phänomens der elektrischen Leitfähigkeit	X		

Tab. 34: Aufgabenauswahl für die Clusteranalyse.

(a) Faktorenanalyse

Mit den zehn ausgewählten Aufgaben wird zunächst eine Faktorenanalyse durchgeführt, als Extraktionsmethode wird die Hauptkomponentenanalyse, als Rotationsverfahren die Varimaxrotation gewählt. Das Ergebnis der Faktorenanalyse zeigt Tabelle 35, dabei wird ersichtlich, dass sich vier Faktoren ergeben, die im Folgenden näher beschrieben werden.

Aufgabe	Faktor			
	1	2	3	4
I.3 „Fragen formulieren“			.74	
I.4 „Versuch zum Gas auffangen“	.66			
I.5 „Reaktionsgleichung“	.72			
II.1 „Kurve beschreiben“			.70	
„Kurve interpretieren“		.78		
II.2 „Fehlersuche“		.56		
II.3 „Kaugummiwirkung“				.89
III.1 „basische Stoffe“	.64			
III.2 „pH-Kurven“		.45		
III.6 „Leitfähigkeit“				.52

Tab. 35: Rotierte Komponentenmatrix mit vier Faktoren; dargestellt sind alle Ladungen > .40.

(b) Beschreibung der Faktoren

Aufgaben, die auf dem ersten Faktor laden

Die Aufgaben „Reaktionsgleichung“, „Versuch zum Gas auffangen“ und „basische Stoffe“ laden gemeinsam auf dem ersten Faktor. Diese drei Aufgaben sind dadurch gekennzeichnet, dass sie typische Aspekte und Anforderungen aus dem Chemieunterricht thematisieren. Sie erfordern das Aufstellen einer Reaktionsgleichung, das Zeichnen einer Versuchsanordnung und das Erkennen und „Übersetzen“ von Symbolformeln verschiedener Stoffe.

Aus dem Lehrerfragebogen ist ersichtlich, dass alle Aufgaben, die auf diesem Faktor laden, generell für z. T. deutlich über die Hälfte der Lehrkräfte eine große Bedeutung für ihren Unterricht haben und dass in mindestens drei Viertel der teilnehmenden Klassen die Inhalte und Anforderungen im Unterricht behandelt wurden. Das lässt sich für die übrigen drei Faktoren nicht in dem Maße feststellen. Zentral für diese Aufgaben ist, dass mit für die Chemie typischen, symbolhaften Repräsentationen (Reaktionsgleichungen und Symbolformeln, Versuchszeichnungen für einen Modellversuch) gearbeitet wird. Dieser Faktor wird im Folgenden als „Schulchemie“ bezeichnet.

Aufgaben, die auf dem zweiten Faktor laden

Auf dem zweiten Faktor laden die Aufgaben „Kurve interpretieren“, „Fehlersuche“ und „pH-Kurven“. In allen drei Aufgaben besteht der Aufgabenstimulus aus einer graphischen Auftragung des pH-Wertes. Bei allen Aufgaben ist es notwendig, dass die Schülerinnen und Schüler die Kurvenverläufe betrachten, um die Aufgabe zu beantworten. Darüber hinaus ist es in allen drei Aufgaben erforderlich, eine qualitative Vorstellung davon zu haben, was der pH-Wert aussagt.

Bei einer Betrachtung der verschiedenen Strukturierungsmerkmale, anhand derer die Testaufgaben klassifiziert wurden, zeigt sich für diese drei Aufgaben, dass es um die Einordnung von Kenntnissen in einen Kontext geht, die ersten beiden Aufgaben im Rahmen des Kontextaspektes Zahnpflegekaugummi, bei der dritten Aufgabe um die Einordnung in den fachchemischen Kontext. Im weiteren Verlauf trägt dieser Faktor die Bezeichnung „pH-Kurven“.

Aufgaben, die auf dem dritten Faktor laden

Die Aufgaben „Fragen formulieren“ und „Kurve beschreiben“, die auf dem dritten Faktor laden, sind dadurch gekennzeichnet, dass zur Beantwortung Text produziert werden muss. Diese beiden Aufgaben haben darüber hinaus gemeinsam, dass sie als eher „unchemisch“ eingeschätzt werden. Über 50 % der unterrichtenden Lehrkräfte geben für diese beiden Aufgaben an, dass sie die benötigten Fähigkeiten im Unterricht eher nicht bereitgestellt wurden, aber zumindest für die Beschreibung von Kurvenverläufen geben ca. 80 % der Lehrkräfte an, dass die Art von Aufgabe für ihren Unterricht generell Bedeutung besitzt. Dieser dritte Faktor wird im Folgenden mit der Bezeichnung „Textproduktion“ umschrieben.

Aufgaben, die auf dem vierten Faktor laden

Die Aufgaben „Leitfähigkeit“ und „Wirkungsweise von Zahnpflegekaugummi“ beschreiben den vierten Faktor, für diese Aufgaben stellen sich die Lehrereinschätzungen sehr unterschiedlich dar. Während die Aufgabe zur Erklärung der Leitfähigkeit für die Lehrkräfte erwartungsgemäß von großer Relevanz ist (alle Lehrkräfte geben an, die benötigten Kenntnisse im Unterricht vermittelt zu haben und bescheinigen der Aufgabe eine generelle Bedeutung für ihren Unterricht), sind die nötigen Fertigkeiten und Kenntnisse zur Beurteilung der Aussagen zur Wirkungsweise von Zahnpflegekaugummi nur von einer Lehrkraft vermittelt worden. Insgesamt schätzen aber ca. 70 % der Lehrkräfte diese Art von Aufgabe in ihrer Unterrichtsbedeutung eher hoch ein. Beide Aufgaben zeichnen sich dadurch aus, dass sie modellhafte Vorstellungen zu Vorgängen in der Wirklichkeit thematisieren. Um die Aufgabe „Wirkungsweise von Zahnpflegekaugummi“ erfolgreich zu beantworten, müssen sich die Schülerinnen und Schüler modellhafte Vorstellungen von den Vorgängen im Mund machen, die ablaufen, wenn sie ein Kaugummi kauen. Für die erfolgreiche Bearbeitung der Aufgabe „Leitfähigkeit“ müssen Teilchenbetrachtungen für eine saure Lösung und destilliertes Wasser vorgenommen werden. Aufgaben, die auf diesem Faktor laden, werden im Folgenden unter der Bezeichnung „modellhafte Vorstellung“ zusammengefasst.

(c) Clusterzentrenanalyse

Anhand dieser vier Faktoren wird eine Clusterzentrenanalyse durchgeführt, um Aussagen über Schülergruppen treffen zu können, die sich anhand der Aufgaben, die die vier Faktoren beschreiben, bilden lassen. Da die Anzahl der Cluster bei diesem Verfahren vorgegeben werden muss, wird im Folgenden beschrieben, wie viele Cluster erwartet werden. Es wird davon ausgegangen, dass sich eine Gruppe finden lässt, die die Aufgaben aller vier Faktoren erfolgreich bearbeitet (Gruppe der leistungsstärkeren Schülerinnen und Schülern), darüber hinaus wird sich eine Gruppe finden lassen, die sich insgesamt aus den leistungsschwächeren Probanden zusammensetzt. Zudem werden Gruppen erwartet, die gute Leistungen hinsichtlich eines Faktors zeigen, während sie auffallende Schwächen in Hinblick auf die anderen Faktoren haben. Insgesamt wird daher davon ausgegangen, dass sich die Stichprobe durch vier bis fünf Cluster erfassen und beschreiben lässt. Es zeigt sich, dass die Stichprobe am besten durch die Annahme von vier Clustern erfasst werden kann. Die entsprechenden Werte für die Berechnungen der Clusterzentren für die Annahme von vier Clustern können den folgenden Tabellen 36a und 36b entnommen werden (Werte für die Annahme von fünf Clustern finden sich im Anhang I, Tab. 53a und 53b, S. 212).

Faktor	Cluster			
	1	2	3	4
„Schulchemie“	.35	- .75	- .65	.97
„pH-Kurven“	.24	.81	- 1.10	.10
„Textproduktion“	.66	.34	- .33	- .67
„modellhafte Vorstellung“	- .94	.69	- .04	.45

Tab. 36a: Clusterzentrenanalyse unter Annahme von vier Clustern, positive Werte sind hervorgehoben.

Cluster	Probanden
1	94
2	84
3	90
4	91

Tab. 36b: Anzahl der Fälle in jedem Cluster, N = 359.

Je größer der angegebene Wert der Clusterzugehörigkeit ist, desto erfolgreicher haben die Schülerinnen und Schüler, die diesem Cluster angehören, die jeweiligen Aufgaben gelöst. Es zeigt sich, dass sich die Cluster am besten darüber beschreiben lassen, mit welchen Aufgaben, die auf einem der vier Faktoren laden, die Lernenden Probleme haben, d. h. welche sie im Vergleich zu den anderen Gruppen nicht erfolgreich lösen können (negativer Wert für die Clusterzugehörigkeit).

(d) Analyse der Cluster: Beschreibung des Antwortverhaltens

Probanden, die dem ersten Cluster zugeordnet werden

Schülerinnen und Schüler, die dem ersten Cluster angehören, sind dadurch gekennzeichnet, dass sich bei den Aufgaben der Faktoren „Schulchemie“, „pH-Kurven“ und „Textproduktion“ positive Clusterzuordnungen finden lassen. Die Aufgaben „Textproduktion“ löst diese Gruppe am erfolgreichsten, während bei den Aufgaben „Schulchemie“ die vierte Gruppe und bei den Aufgaben „pH-Kurven“ die zweite Gruppe am besser abschneiden als diese erste Gruppe. Am schlechtesten von allen Gruppen schneidet diese erste Gruppe bei den Aufgaben „modellhafte Vorstellungen“ ab. Dieser Gruppe werden 94 Probanden und damit etwa ein Viertel (26,2 %) der Gesamtstichprobe zugeordnet, dieser Gruppe gehören die meisten Teilnehmer der Studie an.

Probanden, die dem zweiten Cluster zugeordnet werden

Die zweite Gruppe setzt sich aus 84 Schülerinnen und Schüler (23,4 % der Stichprobe) zusammen. Die Probanden dieses Clusters lösen Aufgaben der Komponente „Schulchemie“ von allen Gruppen am schlechtesten, die Aufgaben „pH-Kurven“ und „modellhafte Vorstellungen“ am besten, bei den Aufgaben „Textproduktion“ ist diese Gruppe die zweiterfolgreichste.

Probanden, die dem dritten Cluster zugeordnet werden

Für die Schülerinnen und Schüler, aus denen sich die dritte Gruppe zusammensetzt, sind für alle vier Faktoren negative Werte feststellbar. Am schlechtesten von allen Gruppen werden die Aufgaben „pH-Werte“ bearbeitet. Diese Gruppen beschreibt 25,1 % der Gesamtstichprobe (90 Probanden).

Probanden, die dem vierten Cluster zugeordnet werden

Diese letzte Probandengruppe zeichnet sich dadurch aus, dass sie die Aufgaben „Schulchemie“ sehr erfolgreich bearbeitet. Ebenfalls positive Werte lassen sich für die Aufgaben der Komponenten „pH-Kurven“ und „modellhafte Vorstellungen“ finden. Ausgesprochen wenig erfolgreich werden von dieser Gruppe die Aufgaben bearbeitet, die der Komponente Textproduktion zugeordnet werden. Diese Gruppe setzt sich aus 91 Schülerinnen und Schülern zusammen und beschreibt damit 25,3 % der Gesamtstichprobe.

(e) Verteilung der Clusterzugehörigkeit verschiedener Stichprobengruppen

Im Folgenden wird aufgezeigt, wie sich die teilnehmenden Klassen hinsichtlich der Anteile der Clusterzugehörigkeit der Lernenden unterscheiden (Abb. 12a und 12b). Darüber hinaus werden

die Häufigkeitsverteilung auf die vier Cluster ermittelt, mit der sich einerseits männliche und weibliche Probanden (Abb. 13) und andererseits die Untersuchungsgruppen mit und ohne „*Chemie-im-Kontext*-Erfahrung“ (Teilgruppe der Gesamtstichprobe, die aus einem Bundesland stammt; Abb. 14) beschreiben lassen.

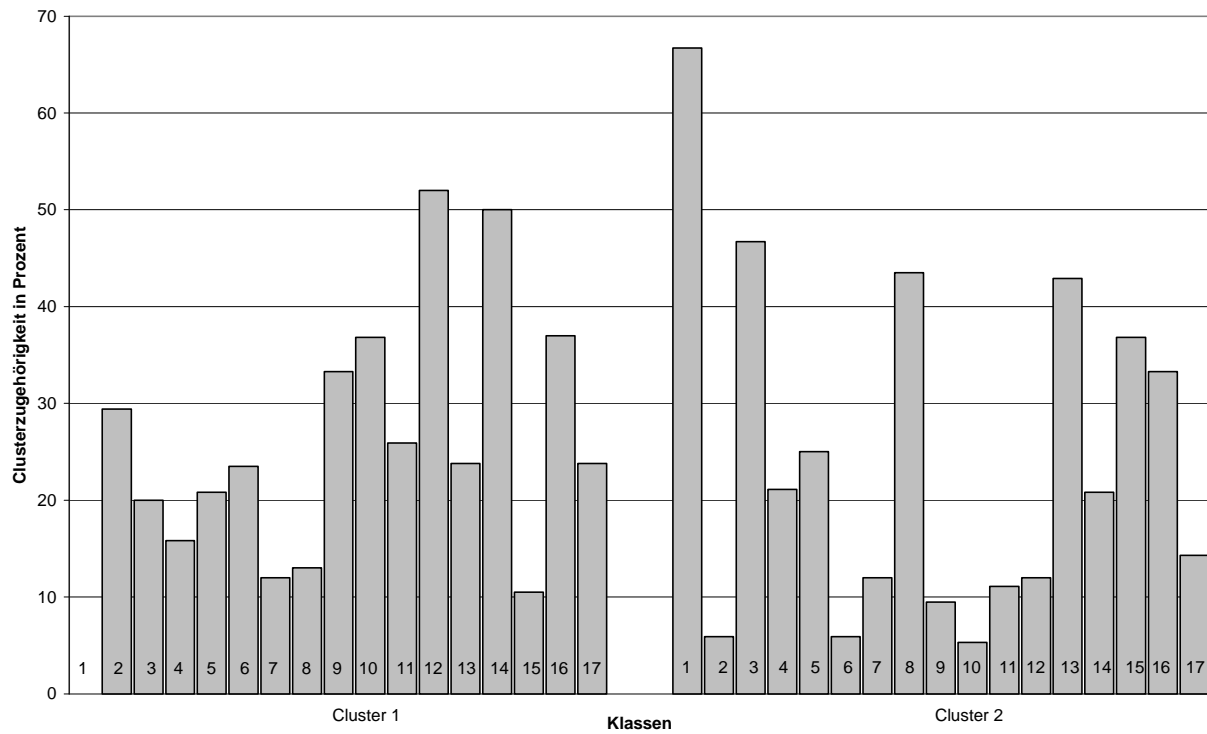


Abb. 12a: Zugehörigkeit zu den Clustern 1 und 2, angegeben der prozentuale Anteil der Klasse 1 – 17; N = 359, Schülerzahlen der einzelnen Klassen können Tab. 22, S. 109 entnommen werden.

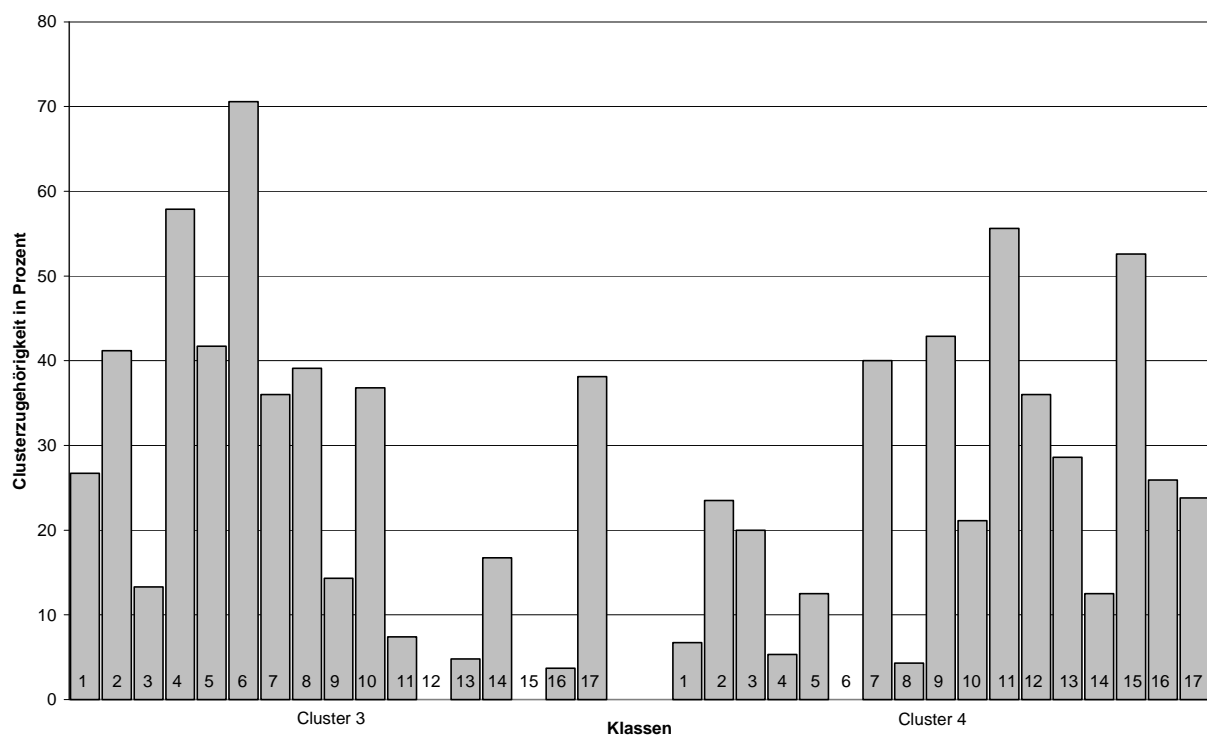


Abb. 12b: Zugehörigkeit zu den Clustern 3 und 4, angegeben der prozentuale Anteil der Klasse 1 – 17; N = 359, Schülerzahlen der einzelnen Klassen können Tab. 22, S. 109 entnommen werden.

Beide Abbildungen geben für jeden der vier Cluster an, welcher Prozentsatz von Schülerinnen und Schülern der jeweiligen Klasse dieser Gruppe zuzuordnen sind. Aufgrund der großen Menge an Informationen, die sich den Abbildungen entnehmen lassen, werden hier nur besonders auffällige Aspekte genannt. Interessant sind die Gruppen, denen aus einzelnen Klassen keine oder nur sehr wenige Schüler bzw. auffallend viele zugewiesen werden: z. B. findet sich im ersten Cluster kein Proband aus Klasse 1 wieder, aber etwa die Hälfte aller Lernenden der Klassen 12 und 14. Deutlich über die Hälfte der Lernenden aus Klasse 1 ist nach ihrem Antwortverhalten dem zweiten Cluster zuzuordnen, mit unter 10 % sind Schülerinnen und Schüler der Klassen 2, 6 und 10 vertreten. Während im dritten Cluster keine Lernenden aus den Klassen 12 und 15 und nur ein geringer Prozentsatz (weniger als 10 %) aus den Klassen 11, 13 und 16 vertreten sind, finden sich die Klassen 4 und 6 mit über der Hälfte der Lerngruppe hier wieder. Dem vierten Cluster werden nur geringe Anteile der Lerngruppen aus den Klassen 1, 4, 6 und 8 zugewiesen, während sich die Klassen 11 und 15 mit über 50 % ihrer Schülerinnen und Schüler dort wieder finden. Betrachtet man die Häufigkeitsverteilung auf die vier Cluster innerhalb der Klassen, zeigt sich ein sehr heterogenes Bild: In keiner Klasse sind alle Cluster etwa in gleichem Maße vertreten, sondern es gibt immer mindestens eine Clusterzugehörigkeit, die nach oben oder nach unten stark aus dem Rahmen fällt.

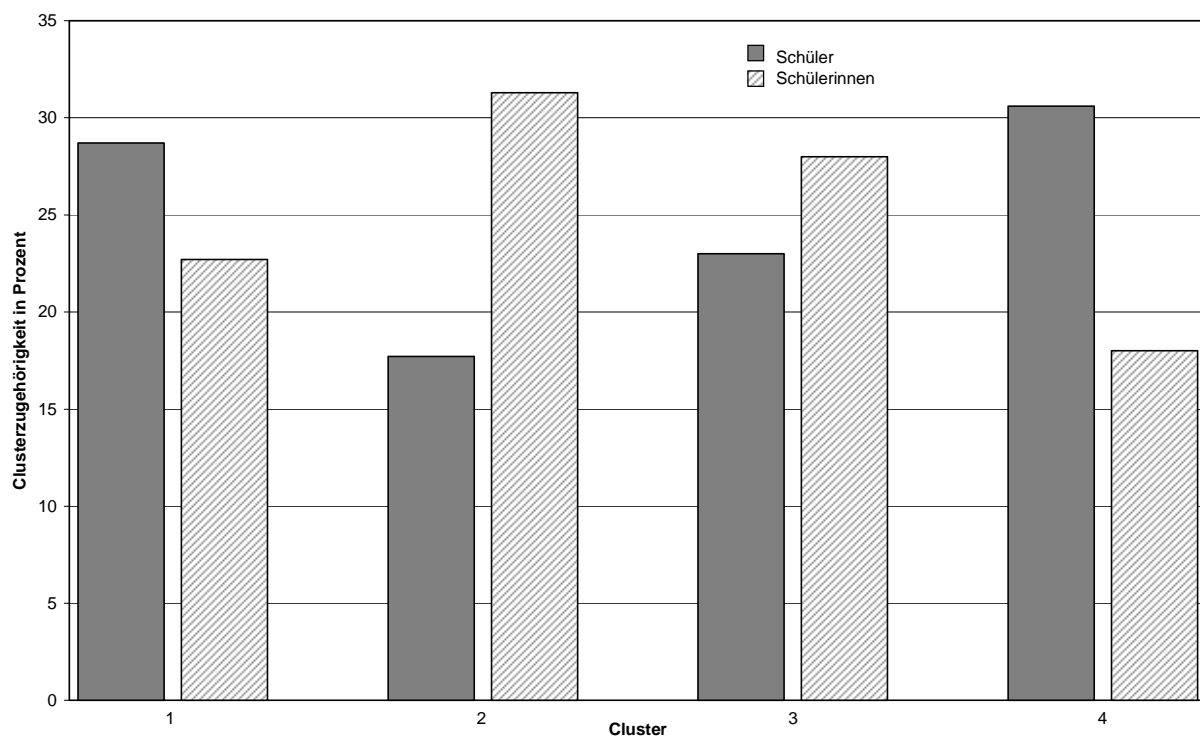


Abb. 13: Clusterzugehörigkeit von Schülern und Schülerinnen im Vergleich; N = 359, m = 209, w = 150.

Bei Betrachtung der geschlechterspezifischen Verteilung auf die vier Gruppen, fällt auf, dass Schülerinnen etwa doppelt so häufig in dem zweiten Cluster zu finden sind wie Schüler. Probanden, die diesem Cluster zugeordnet werden, lösen die Aufgaben, die unter dem Begriff „Schulchemie“ zusammengefasst sind wenig erfolgreich. Im Gegensatz dazu sind sie im vierten Cluster deutlich weniger vertreten als männliche Probanden. Lernende, die diesem Cluster zugeordnet sind, lösen Aufgaben, die eine Textproduktion erfordern, am schlechtesten. Insgesamt fällt auf, dass die Verteilung der Probanden homogener ist als bei Betrachtung einzelner Klassen, die Anteile der zugeordneten Gruppen variieren nicht so stark wie die Verteilung in einzelnen Klassen (Schüler: 17,7 – 30,6 %; Schülerinnen: 18,0 – 31,3 %).

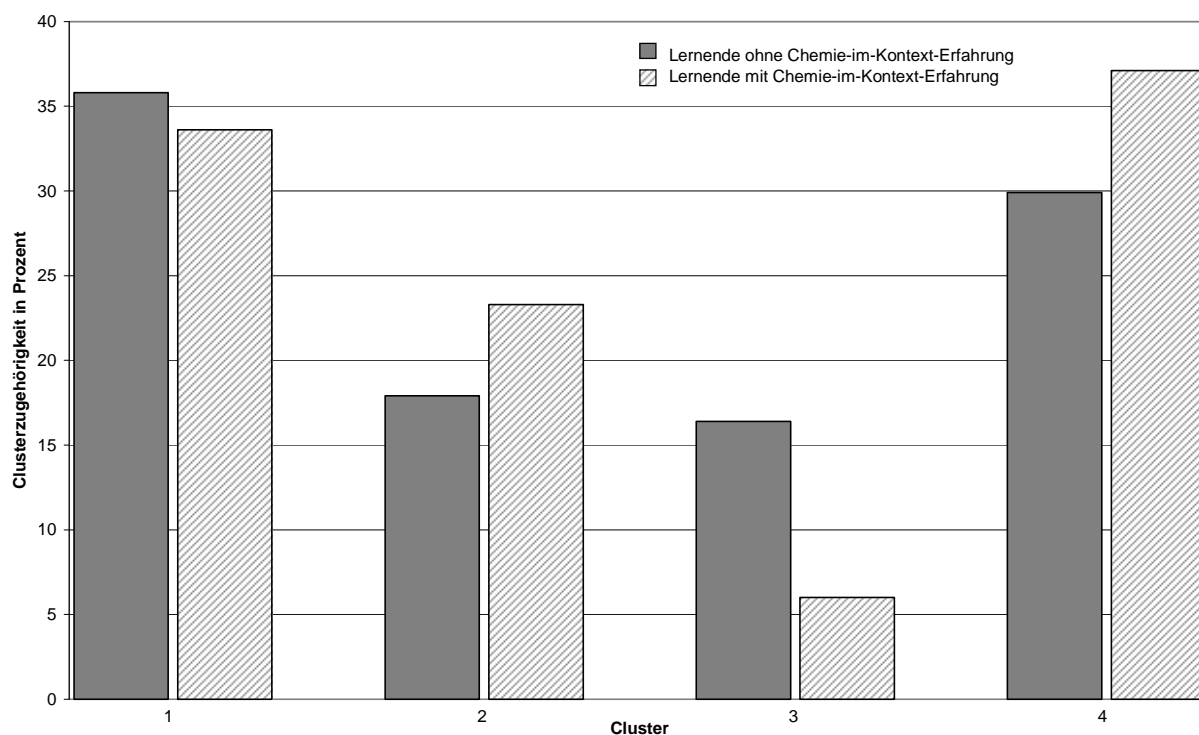


Abb. 14: Clusterzugehörigkeit von Lernenden mit und ohne *Chemie-im-Kontext*-Erfahrung im Vergleich; N = 183, Lernende ohne *Chemie-im-Kontext*-Erfahrung = 67, Lernende mit *Chemie-im-Kontext*-Erfahrung = 116.

Abschließend wird nur eine Teilgruppe der gesamten Stichprobe betrachtet: Die Probanden, die aus einem Bundesland stammen, werden auf ihre Verteilung auf die vier Cluster hin analysiert. Dabei wird die Gruppe der Lernenden, die Erfahrung mit *Chemie im Kontext* hat, mit der Gruppe, denen die Konzeption nicht vertraut ist, verglichen. Die Lernenden mit Konzeptionserfahrung sind nur mit einem auffallend geringen Prozentsatz in dem dritten Cluster vertreten; die Probanden dieses Cluster zeigen in allen Aufgabenbereichen wenig Erfolg. Dagegen ist ein großer Teil dieser Probandengruppe im vierten Cluster zu finden; die Probanden lösen alle Aufga-

ben erfolgreich mit Ausnahme der Aufgaben des Bereichs „Textproduktion“. Bei dieser Probandengruppe variiert die Anteilsgröße in den einzelnen Clustern stärker als bei der Probandengruppe ohne Konzeptionserfahrung und bei Betrachtung der Geschlechtergruppen (Lernenden ohne Konzeptionserfahrung: 16,4 – 35,8 %; Lernende mit Konzeptionserfahrung: 6,0 – 37,1 %). In beiden der hier betrachteten Probandengruppen sind ein Großteil der Schülerinnen und Schüler in den Clustern 1 (defizitärer Aufgabenbereich: „modellhafte Vorstellung“) und 4 (defizitärer Aufgabenbereich: „Textproduktion“) vertreten.

6.4.3.4 Zwischendiskussion 3

Leistungen der Schülerinnen und Schüler

Anhand des Fragebogens können Klassenleistungen differenziert erfasst werden. Während sich in Hinblick auf das Abschneiden im Gesamtttest (bei Betrachtung von 16 der insgesamt enthaltenen 19 Aufgaben) keine signifikanten Unterschiede zwischen den Untersuchungsgruppen, gebildet aufgrund ihres Erfahrungsstand mit der Konzeption *Chemie im Kontext*, feststellen lassen, zeigt eine einfaktorielle Varianzanalyse, dass sich z. T. deutliche Unterschiede zwischen den einzelnen Lerngruppen zeigen.

Dies ist dadurch zu erklären, dass sich die Bedingungen in den einzelnen Klassen sehr inhomogen präsentieren (vgl. Tab. 22, S. 109). Das auffallend schlechte Abschneiden der Klassen 4 und 8 lässt sich durch die geringe Stundenzahl, mit der sich die Klassen jeweils mit dem Thema beschäftigt haben, erklären (Klasse 4: vier Unterrichtsstunden, Klasse 8: drei Unterrichtsstunden). Entsprechend kann das sehr gute Abschneiden der Klasse 16 zumindest zum Teil mit der Dauer der Einheit von 20 Unterrichtsstunden begründet werden. Darüber hinaus ist festzuhalten, dass auch die Streuung innerhalb der Klassen sehr unterschiedlich ist: V. a. Klassen, die im Gesamtabschneiden unter dem Durchschnitt liegen, zeigen innerhalb der Lerngruppe eine größere Streuung als Lerngruppen, die sich im Vergleich als eher leistungsstark präsentieren. Eine Erklärung kann auch in diesem Zusammenhang in der möglicherweise besseren Vorbereitung auf den Test liegen.

Bei Betrachtung der Lehrereinschätzung hinsichtlich der im Unterricht vermittelten Kenntnisse und Fähigkeiten zeigt sich, dass in den drei bzw. vier Unterrichtsstunden der Einheit in den Klassen 4 und 8 natürlich nicht viele Inhalte zu diesem Thema vermittelt wurden, was wiederum das schlechte Abschneiden dieser beiden Klassen insgesamt erklärt. Nach Aussage der Lehrkraft der Klasse 6 sind verhältnismäßig viele Inhalte in der Unterrichtseinheit, die 15 Stunde umfasste, aufgegriffen worden, in dieser Klasse kann ein Grund für das schlechte Abschneiden der Erhe-

bungszeitpunkt darstellen, der direkt vor den Sommerferien nach Abschluss der Zeugniskonferenzen lag. Die Tatsache, dass sich im Abschneiden der Gruppen mit und ohne Unterricht nach *Chemie im Kontext* im Test keine Unterschiede zeigen, lässt sich ggf. damit erklären, dass einige *Chemie-im-Kontext*-Klassen zwar bereits in mehreren Kontexteinheiten im Sinne der Konzeption, aber eben nicht kontinuierlich nach *Chemie im Kontext* unterrichtet wurden. Zudem zeigt die Auswertung des Lehrerfragebogens hinsichtlich der Einschätzung der Lehrkräfte zu Bedeutung der Aufgabenarten für ihren Unterricht generell, dass die Aufgaben, die gerade nicht chemische Fachinhalte in den Vordergrund stellen, sondern in ihren Anforderungen eher den weiteren Zielen der Konzeption zuzuordnen sind, für ihren Unterricht auch eine untergeordnete oder gar keine Bedeutung haben.

In welchem Maße die Lehrkräfte in ihrem Unterricht nach *Chemie im Kontext* die Ziele und Vorgaben der Konzeption tatsächlich realisieren, kann im Rahmen dieser Untersuchung nicht beurteilt und zur Klärung herangezogen werden. Um Aussagen darüber treffen zu können, müsste der Unterricht detailliert systematisch analysiert werden. Möglicherweise lassen sich aus diesen Ergebnissen Hinweise ablesen, dass die konzeptionelle Schwerpunktverschiebung vom traditionellen, an der Fachsystematik orientierten Chemieunterricht hin zu einem auf wenigen Fachkonzepten basierenden, ganzheitlichen Ansatz in vielen Klassen noch nicht (in größerem Maße) stattgefunden hat. Weiteren Aufschluss zu Unterschieden zwischen den Klassen gibt im Folgenden die Analyse des Antwortverhaltens der Lernenden.

Analyse des Antwortverhaltens

Bei Annahme von vier Clustern lassen sich die Besonderheiten im Antwortverhalten der zugeordneten Probanden wie folgt erklären: Insgesamt lässt sich, wie erwartet, eine Gruppe als eher leistungsschwächere Schülerinnen und Schüler beschreiben, sie erlangen für alle vier Bereiche, nach denen die Aufgaben untergliedert sind, negative Clusterwerte, haben demnach mit allen Aufgabenbereichen Schwierigkeiten. Entgegen der Annahme, es ließe sich entsprechend dazu eine Gruppe mit besonders leistungsstarken Lernenden finden, die alle Aufgaben erfolgreich bearbeiten, zeigt sich, dass alle restlichen drei Gruppen einen Aufgabenbereich haben, den sie weniger erfolgreich bearbeiten (Clusterwerte im negativen Bereich) als die übrigen beiden Gruppen. Es handelt sich bei diesen drei Gruppen um „Könner“-Gruppen, die ein Defizit in einer bestimmten Form der Aufgabenanforderung haben. Die Probanden, die dem ersten Cluster zugeordnet sind, haben große Schwierigkeiten, Aufgaben zu bearbeiten, in denen sie modellhafte Vorstellungen zur Klärung heranziehen müssen. Die Probanden des zweiten Clusters schneiden im Aufgabebereich der eher klassischen „Schulchemie“ schlecht ab, während die Schülerinnen

und Schüler, die sich im vierten Cluster wiederfinden, große Schwierigkeiten mit dem Aufgabenbereich haben, der Textproduktion erfordert. Dieser Aufgabenbereich lässt sich anhand der Ergebnisse des Lehrerfragebogens als der Bereich mit dem nach Aussage der teilnehmenden Lehrkräfte geringsten Bezug zum Chemieunterricht klassifizieren. Die Analyse zeigt, dass keine der Gruppen mit Ausnahme der Gruppe, die dem dritten Cluster zugeordnet ist (Cluster 3 fasst die eher leistungsschwächeren Lernenden zusammen), negative Werte im Bereich der Aufgaben mit grafischen Darstellung (pH-Wert-Kurven) hat. Das könnte ein Hinweis darauf sein, dass grafische Darstellungen den Lösungserfolg einer Aufgabe erhöhen (vgl. Prenzel et al. 2002b).

Die prozentuale Zusammensetzung der Klassen nach ihrer Clusterzugehörigkeit bestätigt die Beschreibung des dritten Clusters: Die Tatsache, dass sich hier gehäuft hohe Anteile der Lerngruppen wiederfindet, deckt sich mit dem Ergebnis, dass diese Klassen auch in der Gesamtbetrachtung eher schlecht abschneiden (z. B. Klasse 2, 4, 5 und 6). Die Probanden, die diesem Cluster zugeordnet sind, zeichnen sich durch wenig erfolgreiches Bearbeiten aller vier Aufgabenbereiche aus. Bestätigt wird dieses Ergebnis zusätzlich dadurch, dass diesem Cluster eher keine oder nur sehr wenige Schülerinnen und Schüler der Klassen zugewiesen sind, die in der Gesamtbetrachtung des Tests gut abschneiden (z. B. Klasse 11, 12, 13, 15 und 16). Probanden aus diesen letztgenannten Klassen sind zu einem großen Teil im vierten Cluster zu finden, es kann davon ausgegangen werden, dass sich in diesem Cluster die leistungsstärkeren Schülerinnen und Schüler wieder finden (Klasse 11, 12 und 15), zumal Probanden aus den bereits genannten, im Gesamttest eher unterdurchschnittlich abschneidenden Klassen in diesem Cluster nur sehr gering vertreten sind (Klasse 4, 5 und 6). Die Schülerinnen und Schüler, die diesem Cluster zugeordnet sind, bearbeiten alle untersuchten Aufgabenbereiche erfolgreich mit Ausnahme der Aufgaben, die unter dem Schwerpunkt „Textproduktion“ zusammengefasst sind. Diese Aufgaben könnten von den Lernenden, v. a. von denen, die fachlich wenige Probleme haben, als eher „unchemisch“ und als weniger wichtig wahrgenommen werden. Das könnte ein Grund dafür sein, dass sich gerade die Gruppe der leistungsstarken Lernenden mit Textproduktionsaufgaben eher erfolglos auseinandersetzt.

Die nach Geschlechtern getrennte Betrachtung der Clusterzugehörigkeit zeigt, dass weibliche Probanden zu einem deutlich geringeren Prozentsatz in der Probandengruppe vertreten sind, die alle Aufgabenbereiche mit Ausnahme der Aufgaben zur „Textproduktion“ erfolgreich löst. Dieses Ergebnis könnte möglicherweise auf unterschiedliche sprachliche Fähigkeiten von weiblichen und männlichen Probanden zurückzuführen sein.

Bei der abschließenden Betrachtung der Clusterzugehörigkeit der Stichprobenteilgruppe, aufgeliebert nach ihrer Konzeptionserfahrung, lässt sich feststellen, dass die Probanden, die von Pro-

jektlehrkräften nach *Chemie im Kontext* unterrichtet werden, zu einem auffallend geringen Prozentsatz in der Gruppe der leistungsschwachen, zu einem auffallend hohen Prozentsatz in der Gruppe der leistungsstarken Aufgabenbearbeiter vertreten ist. Das könnte dadurch erklärt werden, dass in der Wahrnehmung der unterrichtenden Lehrkräfte vor und während der Durchführung des Tests das Ergebnis dazu dienen kann oder soll, Aussagen über den Erfolg des Projektes bzw. zur Rechtfertigung der Konzeption zu treffen. Diese Annahme kann dazu geführt haben, dass für die Lehrkräfte zentrale Inhalte und Konzepte zum Themengebiet des Tests besonders intensiv bearbeitet und geübt wurden.

6.4.3.5 Einzelbetrachtung ausgewählter Aufgaben (Fragestellung 3)

Im Rahmen dieser Fragestellung wird u. a. versucht, durch eine vertiefte inhaltliche Analyse der Antworten zu einzelnen Testaufgaben genauere Erkenntnisse zu den Unterschieden in den Klassenleistungen zu erhalten. Darüber hinaus wird geklärt, wodurch Aufgaben gekennzeichnet sind, die sich in Hinblick auf die erfasste Fragebogendimension nicht berücksichtigt werden können. Zudem zeigt sich, dass sich zwar keine Unterschiede zwischen den Untersuchungsgruppen mit und ohne *Chemie-im-Kontext*-Erfahrung im Gesamttest ausmachen lassen, aber durchaus einzelne Aufgaben unterschiedlich erfolgreich durch die Gruppen bearbeitet werden. Diese Aufgaben erfahren im Rahmen dieser dritten Fragestellung eine differenzierte Betrachtung.

(1) Fragestellung 3.1

Welche Besonderheiten kennzeichnen die Aufgaben, die bei der Betrachtung des Gesamttests aufgrund geringer Trennschärfen unberücksichtigt bleiben?

Drei Aufgaben werden aus der Gesamtbetrachtung ausgeschlossen, da sie zu geringe Trennschärfen in Bezug auf die gemessene Dimension „Aspekte des naturwissenschaftlichen Arbeitens“ aufweisen. Während für die Aufgabe I.1 der überwiegende Teil der Lehrkräfte angeben, dass die Aufgabenanforderung im Unterricht thematisiert worden sind, schätzen sie das für die Aufgaben II.3 und II.4 anders ein, hier ist nur jeweils eine Lehrkraft der Ansicht, die Anforderungen auch im Unterricht behandelt zu haben (vgl. Tab. 26a, S. 114).

Zunächst wird Aufgabe I.1 betrachtet, in der die Probanden aufgefordert sind, aus fünf genannten Substanzen die auszuwählen, mit der einer Übersäuerung des Magens entgegengewirkt werden kann. Diese Aufgabe wird von 355 Probanden bearbeitet, von denen 40,6 % die vollständig richtige Alternative „basisch wirkendes Salz“ wählen, der größere Anteil von 48,4 % sich für die

Alternativantwort „Laugen“ oder die Kombination „Laugen“ und „basische wirkendes Salz“ entscheidet. Auf die restlichen Distraktoren verteilt sich insgesamt nur ein geringer Prozentsatz von 11 %. Mit einer Schwierigkeit von $p = 0,65$ liegt diese Aufgabe im mittleren und damit in Hinblick auf Trennschärfetrachtungen günstigen Schwierigkeitsbereich. Abbildung 15 verdeutlicht das unterschiedliche Lösungsverhalten aufgliedert nach den einzelnen Klassen.

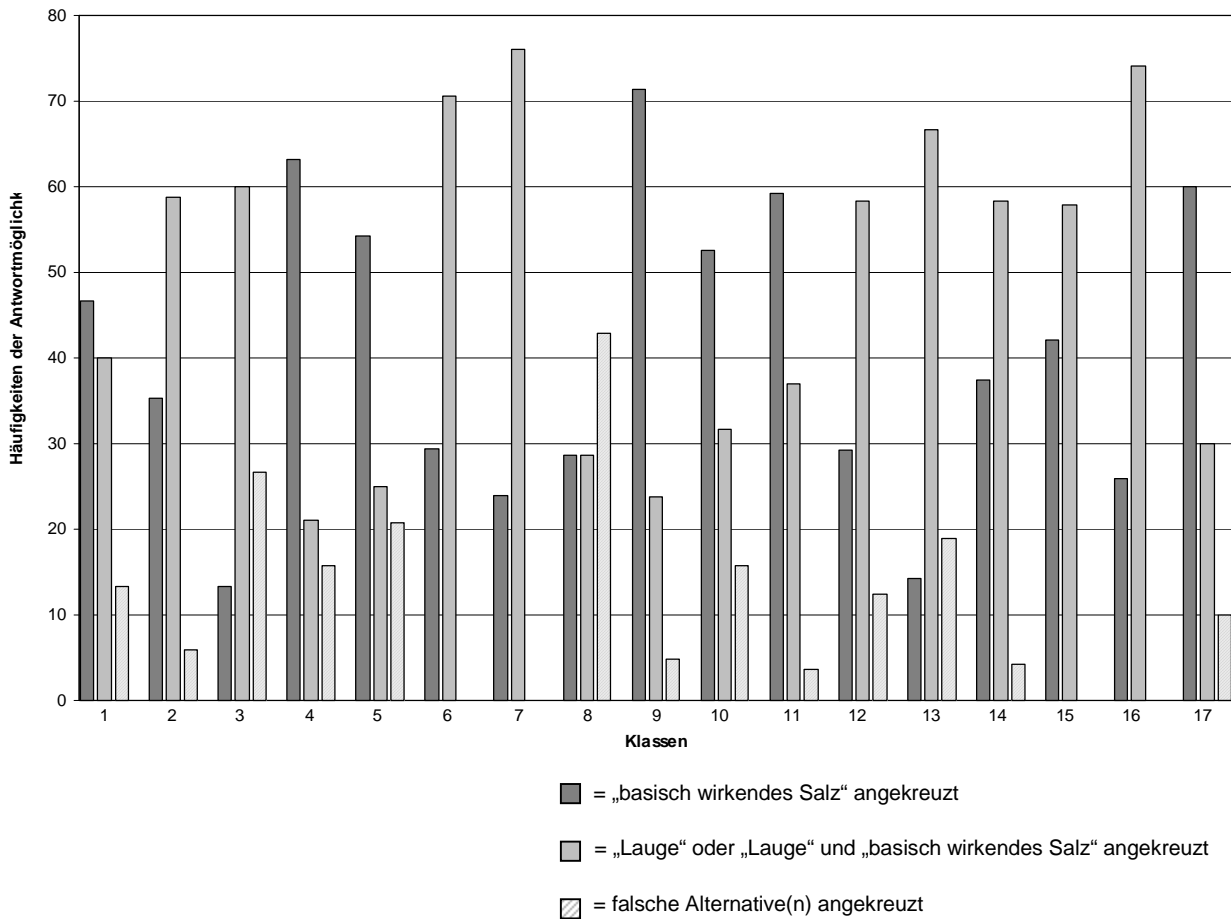


Abb. 15: Aufgabe I.1 – Lösungsverhalten der teilnehmenden Klassen.

Diese Darstellung legt offen, wie unterschiedlich diese Aufgabe in verschiedenen Klassen gelöst wird. Zudem zeigt die Abbildung, dass diese Aufgabe von den Klassen 12, 13, 14, 15 und 16, die insgesamt gut in diesem Test abschneiden, zu einem größeren Prozentsatz die teilrichtige Alternative wählen, während sich in den insgesamt nicht so gut abschneidenden Klassen 4, 5 und 17 der überwiegende Teil der Schülerinnen und Schüler für die korrekte Antwort entscheiden. Damit erklärt sich die geringe Trennschärfe dieser Aufgabe bei der Bestimmung der internen Konsistenz des Gesamttests. In der Bearbeitung der Aufgabe durch die Stichprobenteilgruppe, die nach ihrer Konzeptionserfahrung untergliedert wird, lässt sich kein signifikanter Unterschied feststellen (bei einem betrachteten Signifikanzniveau $< 0,05$, $t_{(138,13)} = - 1,31$, n.s.).

In Aufgaben II.3 sind die Probanden aufgefordert, vier Aussagen zur Wirkungsweise von Zahnpflegekaugummi auf ihren Wahrheitsgehalt hin zu beurteilen. Diese Aufgabe steht nach Aussage der Lehrkräfte kaum mit dem erfahrenen Chemieunterricht in Zusammenhang, wobei aber immerhin fünf von 16 Lehrkräften angeben, dass diese Art von Aufgaben generell für ihren Unterricht von Bedeutung ist. Diese Aufgabe besitzt ebenfalls eine mittlere Schwierigkeit ($p = 0,46$). Diese Aufgabe wird von 351 Probanden bearbeitet, 31,1 % der Schülerinnen und Schüler klassifizieren alle vier Aussagen korrekt, 29,9 % der Teilnehmer der Untersuchung bewerten die Richtigkeit von drei Aussagen korrekt, 39,0 % der Probanden beurteilen den Wahrheitsgehalt von höchstens zwei der vier Aussagen richtig. Die Aussage „Zucker im Kaugummi konserviert die Zähne“ wird von etwa 98 % der Probanden als falsch erkannt, aber nur ca. 65 % der Schülerinnen und Schüler entscheiden sich bei der Aussage „Häufiges Kauen trainiert die Zähne“ richtig, d. h. etwa ein Drittel der Probanden ist der Meinung, dass es möglich ist, durch Kaugummikauen Zähne zu trainieren. Die anderen beiden Aussagen „Durch Kauen angeregter Speichelfluss reinigt die Zähne“ und „Das Kauen eines Kaugummi reinigt die Zähne“ klassifizieren jeweils etwa zwei Drittel der Probanden als richtig (62,2 bzw. 67,4 %).

Auch für diese Aufgaben werden die Häufigkeiten der Antwortmöglichkeiten für die teilnehmenden Klassen betrachtet (Abb. 16).

Im Gegensatz zur zuerst betrachteten Aufgabe zeigt sich bei diesem Testitem, dass in acht der 17 Klassen die Schülerinnen und Schüler am häufigsten höchstens zwei Aussagen richtig klassifizieren. Ein Großteil der Klassen, der im Gesamttest gut abschneidet, löst auch diese Aufgabe erfolgreich (Klassen 3, 11, 13, 15 und 16).

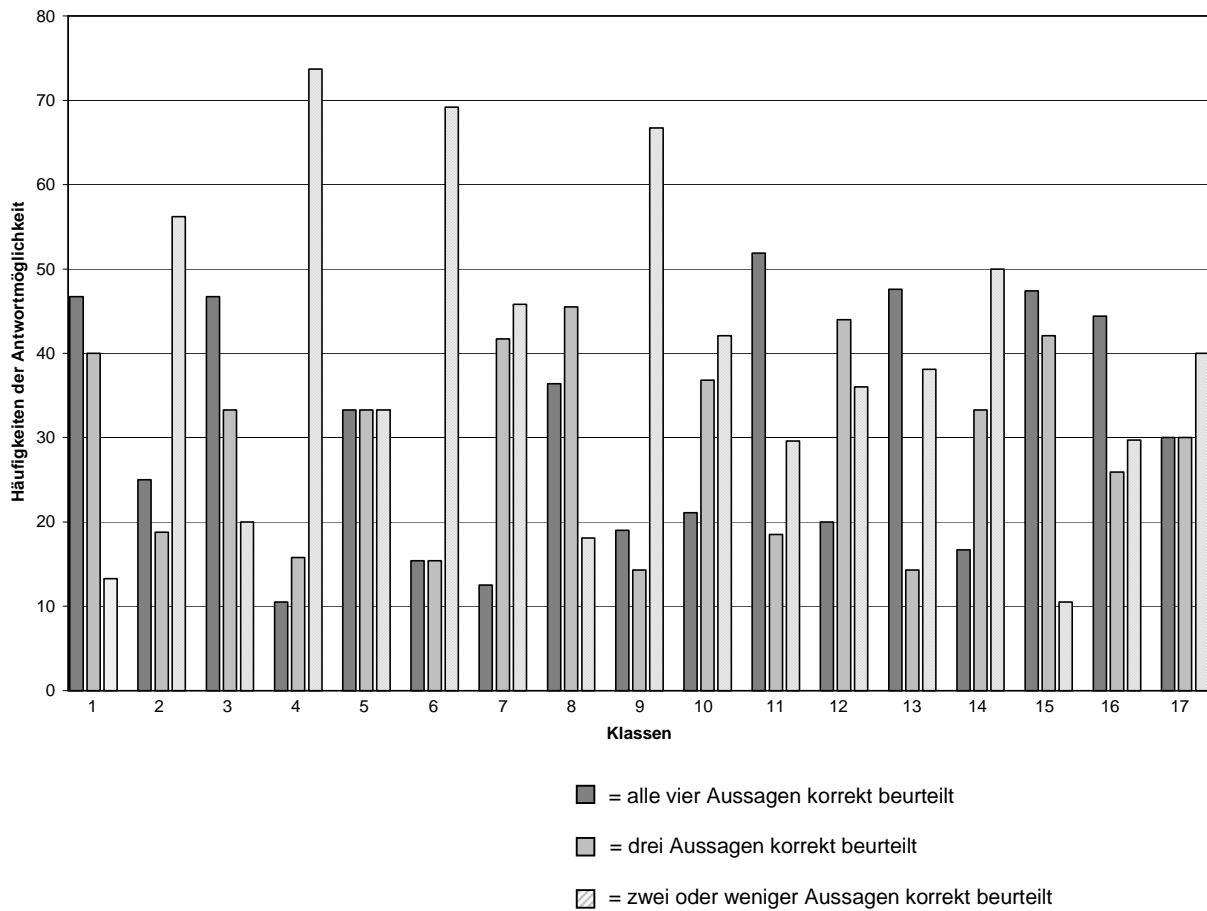


Abb. 16: Aufgabe II.3 – Lösungsverhalten der teilnehmenden Klassen.

Ein Vergleich der aus einem Bundesland stammenden Probandengruppe mit unterschiedlichem Konzeptionshintergrund zeigt keinen signifikanten Unterschied in dem Erfolg der Aufgabenbearbeitung (bei einem betrachteten Signifikanzniveau $< 0,05$, $t_{(135,41)} = 1,34$, n.s.).

Abschließend wird Aufgabe II.4 betrachtet, in der die Schülerinnen und Schüler aufgefordert sind, aus vier gegebenen Aussagen zu einer beschriebenen Studie, die auszuwählen, die eine legitime Ergebnisdarstellung zur Studie beinhaltet.

Insgesamt erkennen 49 % der Probanden, dass die beschriebene Studie keine Aussagen zur Wirksamkeit von Zahnpflegekaugummis zulässt („Diese Studie beweist gar nichts“), die übrigen Probanden kreuzen mehrere Aussagen an oder entscheiden sich für eine falsche Möglichkeit. Es handelt sich bei dieser Aufgabe um ein dichotomes Item, die Itemschwierigkeit liegt bei $p = 0,49$ und damit im mittleren Schwierigkeitsbereich. Bei einer differenzierten Distraktorenbetrachtung wird deutlich, dass nur ein geringer Prozentsatz auf die Aussagen „Diese Studie beweist, dass Kaugummikauen gut für die Zähne ist“ (6,2 %) und „Diese Studie beweist, dass Kaugummikauen schlecht für die Zähne ist“ (1,1 %) entfallen und dass ein hoher Prozentsatz von 45,4 % der Schülerinnen und Schüler der Meinung ist, die Studie zeigt, dass das Kauen eines Kaugummis keine Wirkung auf die Zähne hat. Auch für diese Aufgabe wird eine differenzierte Klassenbetrachtung vorgenommen, die Abbildung 17 zeigt die Häufigkeiten der richtigen und falschen Antworten.

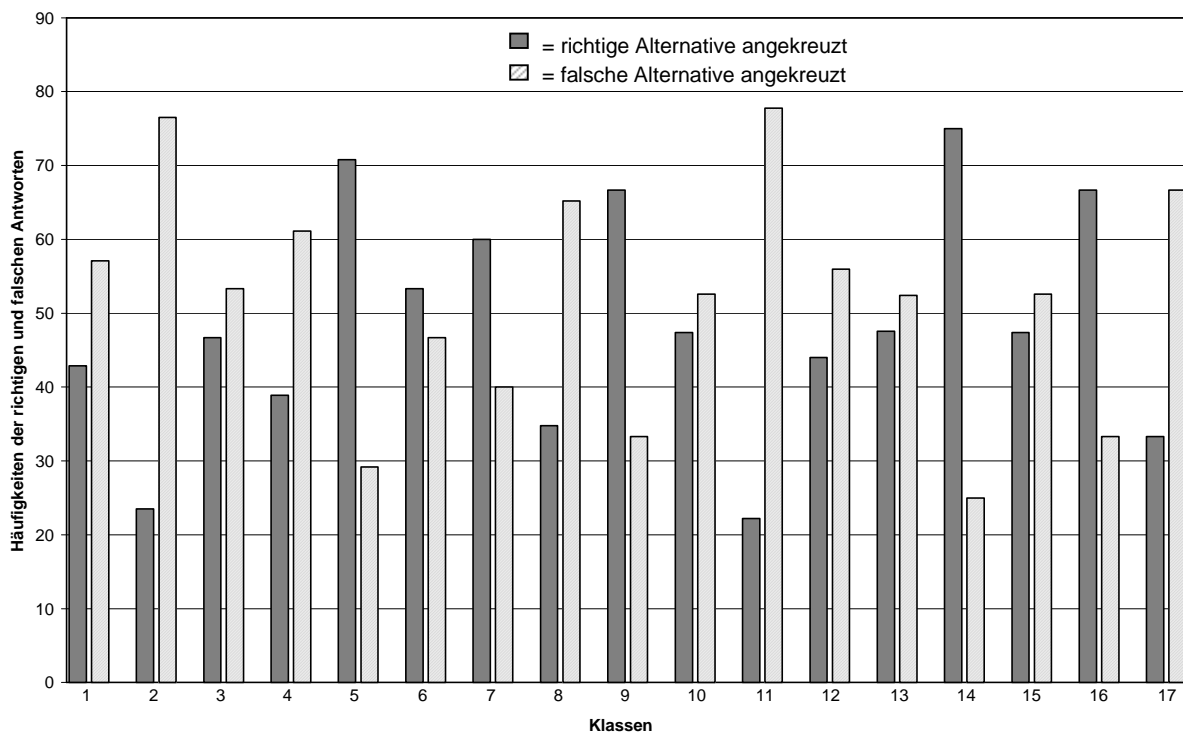


Abb. 17: Aufgabe II.4 – Lösungsverhalten der teilnehmenden Klassen.

Auch hier zeigen sich deutliche Unterschiede zwischen den Klassen, und auch bei dieser Testaufgabe gibt es Klassen, die insgesamt eher in den leistungsschwächeren Bereich gehören, die aber diese Aufgabe auffallend gut lösen (Klasse 5, 6 und 7), während es unter den leistungsstärkeren Klassen Lerngruppen gibt, die bei dieser Aufgaben eher eine falsche Antwortalternative wählen (Klasse 11, 12 und 13).

Wiederum wird untersucht, ob es in dem Erfolg der Bearbeitung dieser Aufgabe einen Unterschied zwischen Probanden ohne und mit *Chemie-im-Kontext*-Erfahrung gibt, und auch in diesem Zusammenhang lässt sich keine Signifikanz feststellen (bei einem betrachteten Signifikanzniveau $< 0,05$, $t_{(140,16)} = - 1,93$, n.s.).

(2) Fragestellung 3.2:

Inwiefern zeigen sich Unterschiede in den Leistungen bei der Bearbeitung einzelner Aufgaben bei Betrachtung zweier Probandengruppen mit unterschiedlichem Erfahrungsstand hinsichtlich der Konzeption Chemie im Kontext bzw. bei einer klassenspezifischer Analyse?

Gruppen mit unterschiedlicher Konzeptionserfahrung

Bei einer vergleichenden Untersuchung einzelner Testitems auf erfolgreiche Bearbeitungen durch die Stichprobenteilgruppen aus einem Bundesland, die sich in ihrer Konzeptionserfahrung unterscheiden, zeigen sich bei sechs der insgesamt 16 betrachteten Items signifikante Unterschiede. Von diesen sechs Testaufgaben werden vier von den Probanden ohne Konzeptionserfahrung signifikant besser gelöst als von der Probandengruppe mit Konzeptionserfahrung. Tabelle 37 gibt die entsprechenden Aufgaben wieder.

Testaufgabe	N (Prob. mit Konzeptionerf.) N (Prob. ohne Konzeptionerf.)	Mittelwerte	Angaben der t-Test-Werte	Signifikanzniveau
I.2	114	1,34	$t_{(152,02)} = - 2,3$	$< .05$
	66	1,61		
I.5b	82	0,41	$t_{(106,82)} = - 2,4$	$< .05$
	60	0,78		
II.1a	113	0,34	$t_{(127,17)} = - 2,4$	$< .05$
	65	0,52		
III.1	113	0,46	$t_{(134,76)} = - 2,4$	$< .05$
	64	0,64		
III.2	116	0,88	$t_{(104,84)} = 2,4$	$< .05$
	66	0,73		
III.5	108	1,16	$t_{(82,61)} = 2,3$	$< .05$
	48	0,77		

Tab. 37: Unterschiede im Bearbeitungserfolg einzelner Aufgaben; betrachtete Stichprobe: Lernenden aus einem Bundesland.

Um zu diesen Ergebnissen weitere Erkenntnisse zu erlangen, werden ausgewählte Aufgaben exemplarisch untersucht.

Zunächst wird die Aufgabe I.2 betrachtet. Diese Aufgabe liegt mit ihrem Schwierigkeitsgrad im mittleren Bereich ($p = 0,64$), die Probanden sind aufgefordert, Kriterien zu nennen, die bei der Entwicklung eines Medikaments von Bedeutung sind. Alle von den Lernenden genannten Kriterien wurden dahingehend inhaltlich analysiert, ob sie eher aus dem chemischen, biologischen oder medizinischen Bereich stammen (z. B. Aussagen zur Verträglichkeit, zu Nebenwirkungen, zur chemischen Reaktion) oder ob Kriterien genannt werden, die sich u. a. eher „kaufmännischen“ oder „alltäglichen“ Überlegungen zuordnen lassen (z. B. Aussagen zum Preis, Geschmack, zur Werbung, Farbe, Verpackung). Probanden mit Konzeptionserfahrung nennen aus diesem zweitgenannten Bereich deutlich weniger Kriterien (39 % der Probanden) gegenüber den Lernenden ohne Konzeptionserfahrung (65 % der Probanden).

Nach Aussage der unterrichtenden Lehrkräfte dieser Klassen hat keine die notwendigen Kenntnisse und Fähigkeiten im Unterricht thematisiert und sie schreiben dieser Aufgabenart für ihren Unterricht keine Bedeutung zu.

Ebenfalls besser schneidet diese Gruppe bei der Berechnungsaufgabe I.5b ab; insgesamt ein größerer Prozentsatz dieser Probandengruppe bearbeitet die Aufgabe (90 % der Lernenden dieser Gruppe im Vergleich zu 70 % der Probanden mit Konzeptionserfahrung). Nach Einschätzung der unterrichtenden Lehrkräfte beider Probandengruppen sind in allen drei Klassen ohne Konzeptionserfahrung die Anforderungen dieser Aufgabe thematisiert worden, in den Klassen mit Konzeptionserfahrung wurden sie in einer von fünf Klassen nicht thematisiert. Im Durchschnitt schätzen die Lehrkräfte beider Probandengruppen die Lösungshäufigkeit ihrer Klassen gleich ein: für Klasse ohne Konzeptionserfahrung geben die Lehrkräfte eine durchschnittliche Lösungshäufigkeit von 43 %, für die Klasse mit Konzeptionserfahrung liegt die Angabe bei 42 %. Die Ergebnisse der klassenspezifischen Betrachtung zeigt folgendes Bild zum Bearbeitungserfolgs: Während sechs der insgesamt acht betrachteten Klasse bei dieser Aufgabe eine Lösungshäufigkeit zwischen 0 und 20 % zeigen (Klasse 9: 0 %, Klasse 10: 15 %; Klasse 11: 7,4 %, Klasse 13: 14 %, Klasse 14: 21 %, Klasse 15: 0 %), fallen eine Klasse ohne *Chemie-im-Kontext*-Erfahrung mit einer Lösungshäufigkeit von 74 % (Klasse 16) und eine Klasse mit *Chemie-im-Kontext*-Erfahrung mit einer Lösungshäufigkeit von immerhin 50 % (Klasse 12) heraus.

Im folgenden Abschnitt wird Aufgabe III.2 betrachtet, in der Schülerinnen und Schüler besser abschneiden, die bereits in mindestens drei Einheiten nach *Chemie im Kontext* unterrichtet wurden. In dieser Aufgabe sind die Probanden aufgefordert, einem beschriebenen Titrationsvorgang den richtigen Kurvenverlauf zur pH-Wert-Änderung zuzuweisen. Diese Aufgabe hat mit $p = 0,72$ einen niedrigen Schwierigkeitsgrad, entsprechend haben auch beide Probandengruppe hohe Lösungshäufigkeiten zu verzeichnen (mit Konzeptionserfahrung 88 %, ohne Konzeptionserfahrung 73 %). Während alle Klassen, die nach *Chemie im Kontext* unterrichtet werden, ausnahmslos hohe Lösungshäufigkeiten bei dieser Aufgabe verzeichnen (81 – 95 %), hat nur eine der drei Klasse, die keine Erfahrung mit der Konzeption haben, mit 96 % eine (Klasse 16) ebenfalls hohe Lösungshäufigkeit, in den anderen beiden Klassen wird diese Aufgabe nur von 46 % (Klasse 9) bzw. 33 % (Klasse 10) der Lernenden korrekt beantwortet. Bei Betrachtung der Ankreuzhäufigkeiten der übrigen Antwortalternativen lässt sich feststellen, dass ein vergleichsweise hohe Prozentsatz die Antwortalternative E (pH-Wert-Verlauf von pH 14 bis 1) wählt (Klasse 9: 19 %, Klasse 10: 44,4 %). Alle drei unterrichtenden Lehrkräfte der Klasse 9, 10 und 16 geben an, dass die Aufgabenanforderungen im Unterricht thematisiert wurden, die durchschnittliche Lösungshäufigkeit nach Einschätzung der Lehrkräfte liegt bei 57 %. Demgegenüber wurden diese Anforderungen in drei der fünf übrigen Klassen nicht bearbeitet (durch die Lehrkräfte geschätzte durchschnittliche Lösungshäufigkeit 46 %).

Klassenspezifische Betrachtungen

Abschließend werden die Schülerbearbeitungen zu Aufgabe III.6 genauer analysiert, da diese Aufgabe einen zentralen Aspekt des Donator-Akzeptor-Basiskonzepts thematisiert. Zudem haben alle unterrichtenden Lehrkräfte die Aufgabeanforderungen nach eigener Einschätzung thematisiert und schätzen daher die Lösungshäufigkeit für ihre Lerngruppe entsprechend hoch ein (durchschnittlich eingeschätzte Lösungshäufigkeit liegt bei 58 %), weitere Daten zur eingeschätzten und tatsächlichen Lösungshäufigkeit ist in Abbildung 18 dargestellt.

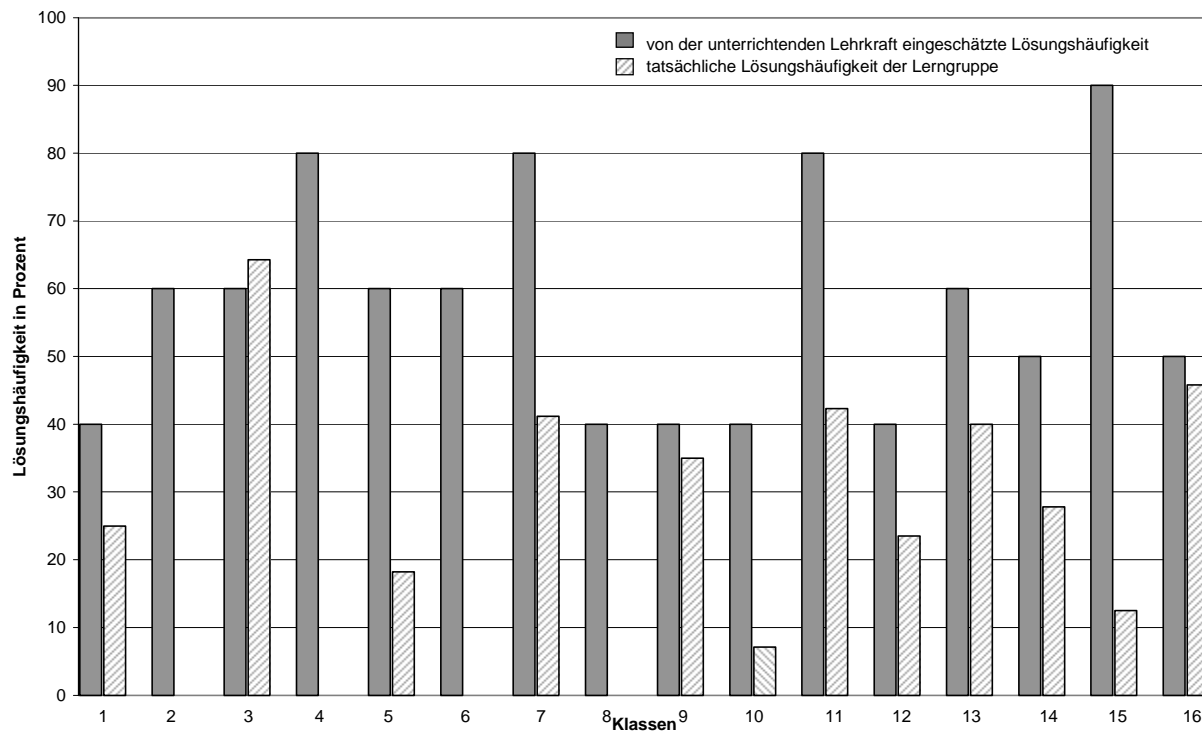


Abb. 18: Von der unterrichtenden Lehrkraft eingeschätzte und tatsächliche Lösungshäufigkeit für Aufgabe III.6 (Erklärung der Leitfähigkeit); in dieser Darstellung werden nur als vollständig richtig gewertet Schülerantworten berücksichtigt, da die Lehrkräfte anhand dieser Antwortvorgabe entsprechend dem Erwartungshorizont ihre Einschätzung gegeben haben; für Klasse 17 liegt keine Einschätzung durch die Lehrkraft vor.

Der Schwierigkeitsgrad der Aufgabe liegt mit $p = 0,43$ im mittleren Bereich, dennoch liefert aus einigen Klassen (Klasse 2, 4, 6 und 8) kein einziger Proband eine vollständig richtige Lösung. In drei Fällen stimmt die Lehrereinschätzung für diese Aufgabe gut mit dem tatsächlichen Ergebnis überein (Klasse 3, 9 und 16), in den anderen Fällen liegt die Lehrereinschätzung z. T. deutlich über den tatsächlichen Leistungen.

Um eine Erklärung für das in einigen Fällen sehr schlechte Abschneiden der Klassen finden zu können, werden die Schülerantworten aus einzelnen Klassen inhaltlich analysiert.

Insgesamt lassen sich verschiedene Konzepte ausmachen, die im Rahmen der Auswertung der Aufgabe als falsch gewertet werden: Ein Teil der Lernenden geht davon aus, dass die Säure bzw. das Wasser geladen sind und daher den Strom leiten (z. B. „Destilliertes Wasser ist neutral und somit nicht positiv oder negativ geladen, saure Lösung ist positiv geladen“, „Da Wasser neutral ist, hat es keine negativen bzw. positiven Teilchen, die mit elektrischem Strom ausgeglichen werden müssen. Saure Lösungen sind unter dem neutralen Wert“).

Eine weitere, häufig auftretende Schülererklärung beruft sich auf das Vorhandensein freier Elektronen oder ein Überschuss an Elektronen (in der sauren Lösung/in der Säure) (z. B. „Wasser = neutral, z. B. HCl = freie Elektronen“, „Die saure Lösung hat mehr freie Elektronen, die

den Strom leiten. Destilliertes Wasser hingegen hat keine freien Elektronen“, „Die Leitfähigkeit ist abhängig von der Anzahl der Elektronen, $\text{H}_2\text{O} = 10$ Elektronen, $\text{H}_2\text{SO}_4 = 50$ Elektronen“).

Ein drittes, häufig auftretendes Erklärungsmuster in den Schülerantworten ist die Existenz von Stoffen, die für die Leitfähigkeit verantwortlich sind (z. B. „Im destillierten Wasser ist kein Eisen vorhanden, so muss in einer sauren Lösung Ähnliches enthalten sein, damit Strom fließt“, „Destilliertes Wasser ist nur Wasser ohne Nährstoffe, saure Lösung ist Flüssigkeit mit Nährstoffen, welche den Strom leiten“, „Im destillierten Wasser befinden sich fast keine „Metalle“ mehr, die den Strom leiten. Die Säure dagegen hat einen Metallanteil, der den Strom leitet“).

In einzelnen Klassen treten die fehlerhaften Erklärungsmuster in den Schüleräußerungen gehäuft auf: In Klasse 10 finden sich z. B. von insgesamt 14 (fünf Probanden haben diese Aufgabenstellung nicht bearbeitet) acht Antworten, die (freie) Elektronen in der sauren Lösung für den Stromfluss verantwortlich machen.

In Klasse 2, in der keine Schüleräußerung als vollständig richtig klassifiziert wird, sind neben 39 % teilrichtigen Lösungen unter den als falsch klassifizierten Antworten alle drei hier vorgestellten Erklärungsmuster vertreten.

Eine vollständig richtige Antwort wird z. B. in Klasse 4 bei keinem Probanden klassifiziert, aber immerhin 42 % der Lernenden geben eine teilrichtige Lösung an, falsche Lösungen lassen sich auf das Konzept von Stoffen in der Lösung, die an sich leitfähig sind wie z. B. Metall, oder fachlich falsche Aussagen wie „Wasser ist kein Dipol“ zurückführen.

6.4.2.6 Zwischendiskussion 4

Aufgaben mit geringer Trennschärfe

Die Tatsache, dass in Aufgabe I.1 von den Probanden nicht explizit gefordert wird, sich für eine Antwortalternative zu entscheiden, scheint bei den Lernenden zu einer Verunsicherung zu führen, so dass sie sich häufig für beide aus chemischer Sicht richtigen Antworten entscheiden und den Schritt zur Beurteilung der praktischen Folgen und damit den Bezug zum tatsächlichen Leben nicht herstellen. Insbesondere bei den leistungstärkeren Probanden ist dieses Antwortverhalten damit zu erklären, dass sie aus chemischer Sicht beide richtigen Antworten erkennen und ankreuzen. Da es sich um eine Multiple-Choice-Aufgaben handelt, sind die Möglichkeiten einer weiteren inhaltlichen Analyse beschränkt. Es kann festgehalten werden, dass die Schülerinnen und Schüler, die sich für die Alternative „Laugen“ oder die Kombination „Laugen“ und „basisch wirkendes Salz“ entscheiden, zur Beantwortung der Aufgabe auf den chemischen Rahmen der Aufgabe zurückgreifen, ohne den kontextlichen Rahmen der Aufgaben in ihre Antwortfindung

einzu beziehen. Wird die Tatsache, dass ein Mensch an der Übersäuerung leidet und daher die Substanz zu sich nehmen können muss, berücksichtigt, schließt sich die Antwort „Lauge“ aus. Die geringe Trennschärfe ist möglicherweise dadurch zu begründen, dass die Schülerinnen und Schüler anhand der Aufgabenstellung nicht aufgefordert waren, sich für eine Antwort zu entscheiden oder ihre Antwort zu begründen, somit könnte die Aufgabenstellung insbesondere leistungsstärkere Probanden verunsichert haben.

Für Aufgabe II.3 und II.4 ist festzuhalten, dass beiden Aufgaben von den Lehrkräften eine geringe Bedeutung für ihren Unterricht bescheinigt wird. Das gilt auch für Projektlehrkräfte, obwohl diese beiden Aufgaben als eher typisch für einen Unterricht nach *Chemie im Kontext* eingestuft werden. Die Aufgaben II.3 und II.4 messen ein Alltagsverständnis hinsichtlich naturwissenschaftlicher Vorgänge und Methoden der Erkenntnisgewinnung. Sie sind ohne fachchemischen Hintergrund lösbar und passen daher vermutlich nicht in die eher „chemieunterricht-konforme“ Dimension des naturwissenschaftlichen Arbeitens, die der Fragebogen erfasst.

Aufgabe II.3 scheint bei den Probanden in erster Linie unvollständiges Alltagswissen zum Kaugummikauen abzurufen. Mit Ausnahme der Aussage zur Konservierung der Zähne durch Zucker klingen die übrigen drei Aussagen für jeweils etwa zwei Drittel der Probanden plausibel.

Bei Aufgabe II.4 könnte so etwas wie der „Glaube an die Wissenschaft“ immerhin etwa Hälfte der Probanden dazu verleiten, eine der falschen Antwortalternativen zu wählen. Während die ersten drei Antwortalternativen sich konkret auf die Auswirkungen auf die Zähne (gut, schlecht, keine Wirkung) beziehen, hat der vierte Distraktor eine andere Qualität, indem er die Gesamtanlage der Studie als Bezugspunkt hat. Für die Lernenden stehen offensichtlich nicht so sehr der Aufbau der Studie, sondern die Ergebnisse im Vordergrund. Daher wird auch, ergebnisorientiert, das Ausmaß der Wirkung beurteilt. Diesen Probanden fehlt die Fähigkeit zur Beurteilung der Aussagekraft.

Probanden mit und ohne Konzeptionserfahrung: Einzelbetrachtung ausgewählter Aufgaben

Es fällt auf, dass gerade bei einer Aufgabe, für die angenommen werden kann, dass sie von Lernenden mit Konzeptionserfahrung besser oder doch zumindest ebenso erfolgreich bearbeitet wird, diese Probandengruppe signifikant schlechter abschneidet als die Probandengruppe ohne Konzeptionserfahrung. Es kann davon ausgegangen werden, dass die Lernenden, die keinerlei Erfahrung mit kontextbasiertem Chemieunterricht haben, unvoreingenommener an diese Aufgabe herangehen als Lernende, die bereits in mehreren Einheiten nach *Chemie im Kontext* unterrichtet wurden. Während die Probandengruppe, die bisher nicht an kontextbasierten Unterricht teilgenommen hat, folglich frei assoziiert und auch ihre Alltagserfahrungen in die Überlegungen

einbezieht, behalten die Lernenden aus *Chemie-im-Kontext*-Klassen stärker Aspekte des Schul- bzw. Chemieunterrichts im Blick und scheinen die Aufgabe nicht vorbehaltlos zu bearbeiten. Das kann der Grund dafür sein, dass sie bei dieser Aufgabe schlechter abschneiden, eben weil sie stärker zielgerichtet an der Lösung arbeiten und damit ihre Antwortmöglichkeiten einschränken. Ursache dafür könnte der bis dahin wahrgenommene Verlauf von Kontexteinheiten sein, die zwar von einer offenen Fragestellung ausgehen, aber letztlich doch chemische Inhalte fokussieren. Lernende, die keine Erfahrung mit einem Unterricht besitzen, der von lebensweltlichen Fragestellungen ausgeht, trauen sich eher, auch „kaufmännische“ Aspekte zu notieren.

Diese Gruppe hat ebenfalls mehr Erfolg bei der Bearbeitung der Aufgabe I.5b, was damit zusammenhängen könnte, dass insgesamt ein größerer Prozentsatz dieser Probandengruppe die Aufgabe überhaupt bearbeitet. Die Ergebnisse der klassenspezifischen Betrachtung des Bearbeitungserfolgs liefern einen Hinweis zur Erklärung des unterschiedlichen Abschneidens: Insbesondere aus Klasse 16, die keine Konzeptionserfahrung hat, lösen drei Viertel der Probanden diese Aufgabe mit Erfolg, in der erfolgreichsten Klasse mit *Chemie-im-Kontext*-Erfahrung bearbeitet die Hälfte der Klasse die Aufgabe richtig. Der Unterschied ist folglich auf einen Klasseneffekt zurückzuführen. In Klasse 16 ist diese Art der Berechnung möglicherweise besonders vertiefend bearbeitet worden oder aber der zeitliche Abstand zwischen Thematisierung dieser Aufgabenanforderung und Testbearbeitung war vergleichsweise klein.

Abschließend wird mit Aufgabe III.2 eine Aufgabe betrachtet, die eine grafische Darstellung im Aufgabenstimulus enthält. Diese Aufgabe wird von der Probandengruppe mit *Chemie-im-Kontext*-Erfahrung erfolgreicher bearbeitet als von der anderen Gruppe. In beiden Fällen wird eine Diskrepanz zwischen der eingeschätzten und der tatsächlichen Lösungshäufigkeit deutlich: Die unterrichtenden Lehrkräfte unterschätzen ihre Lerngruppen hinsichtlich ihrer Fähigkeit, die Aufgabe zu lösen. Zwei Klassen, in denen die Aufgabenanforderung behandelt wurde, ist im Verlauf des Unterrichts offensichtlich ein Missverständnis hinsichtlich der pH-Werte von sauren und basischen Lösungen aufgetreten, denn die Lernenden, die sich für die Antwortalternative E entscheiden, erkennen richtig, dass der pH-Wert sich nicht dem neutralen Bereich annähert, ordnen die pH-Werte 1 und 14 aber nicht korrekt zu. Wiederum kann dieser signifikante Unterschied zwischen den beiden untersuchten Gruppen auf einen Klasseneffekt zurückgeführt werden. Darüber hinaus senkt der grafische Aufgabenstimulus möglicherweise die Aufgabenschwierigkeit in größerem Maße (vgl. Prenzel 2002b) als von den Lehrkräften erwartet.

Insgesamt lässt sich festhalten, dass mit Ausnahme der ersten Aufgabe, das unterschiedliche Abschneiden der beiden untersuchten Probandengruppen eher mit Klasseneffekten als mit über-

geordneten Aspekten, wie eben der unterschiedliche Erfahrungshintergrund hinsichtlich der Konzeption *Chemie im Kontext* erklärt werden kann.

Lehrereinschätzung und Aufgabenbearbeitung

Insbesondere die große Diskrepanz zwischen Lehrereinschätzung und tatsächlicher Leistung der Lerngruppe fällt bei der Betrachtung der Testaufgabe III.6 ins Auge. Die Angaben der Lehrkräfte, dass sie die Aufgabenanforderungen im Unterricht behandelt haben, und die Tatsache, dass diese Art der Aufgabe von allen Lehrkräften für ihren Unterricht generell als bedeutsam eingestuft wird, rechtfertigt ihre Einschätzung hinsichtlich der Lösungshäufigkeit. Die Gründe müssen sicherlich für die einzelnen Klassen gesondert betrachtet werden, aber dennoch lassen sich zwei übergeordnete Erklärungen anführen, die für alle Lerngruppen gleichermaßen gelten: die Stellung der Testaufgabe am Ende des Fragebogens und eine möglicherweise „strengere“ Beurteilung der Schülerantworten durch Außenstehende.

Klassenspezifische Gründe, die hier vermutet werden können, sind z. B. gehäuft auftretende fehlerhafte Erklärungsmuster (z. B. Klasse 10), deren Entstehung bzw. Festigung sich nur in Rücksprache mit der Lehrkraft und mit Blick auf den tatsächlichen Unterricht geklärt werden kann, oder eine geringe Stundenzahl, mit der die Einheit unterrichtet wurde (Klasse 4 und 8).

6.4.3.7 Diskussion

Insgesamt erweist sich der Fragebogen als brauchbares Testinstrument zur Feststellung des Lernstands von Schülerinnen und Schüler. Es ist möglich, vergleichende und differenzierende Aussagen zu Klassenleistungen und dem Abschneiden ausgewählter Probandengruppen zu treffen. Es muss aber festgehalten werden, dass es nicht gelungen ist, Testaufgaben zu entwickeln, mit dem einzelne Kompetenzen oder kognitive Anforderungen isoliert erfasst werden können, vielmehr handelt es sich um ein eindimensionales Instrument, mit dem der Lernstand zur Facette des naturwissenschaftlichen Arbeitens im Themenbereich „Säuren, Basen und Neutralisation“ insgesamt erhoben wird.

Während zwischen den teilnehmenden Klassen deutliche Unterschiede feststellbar sind, ergibt ein Vergleich zweier Gruppen, die danach gebildet sind, ob sie einen Unterricht nach *Chemie im Kontext* erfahren haben oder nicht, keine signifikanten Unterschiede. Mehrere Gründe können dafür angeführt werden: der Effekt der unterrichtenden Lehrkraft ist größer als der der Konzeption, an der sich die Lehrkraft orientiert. Darüber hinaus brauchen Veränderungen Zeit, um Wirkungen zeigen zu können. Sicher können einige kontextorientierte Unterrichtseinheiten, die auf

zwei Schuljahre verteilt sind, keine große Effekte hervorrufen. Um wirklich gesicherte Aussagen zu einem Vergleich von „herkömmlichen“ Unterricht und einem Unterricht nach *Chemie im Kontext* treffen zu können, müsste sich *Chemie im Kontext* als durchgängige Konzeption für den Chemieunterricht etabliert haben. Darüber hinaus bedarf es Kriterien und Verfahren, die es ermöglichen, einen Unterricht als einen Unterricht nach *Chemie im Kontext* zu kennzeichnen, was wiederum eine Systematisierung von Unterricht voraussetzt.

Zur Stichprobe dieser Untersuchung muss festgehalten werden, dass ein Teil der Lehrkräfte, die sich auf das Projekt *Chemie im Kontext* eingelassen haben, den Grundlagen, v. a. der Zurücknahme der Fachsystematik als strukturierendes Element, eher skeptisch gegenüber stehen, da u. a. ein Niveauverlust im Chemieunterricht befürchtet wird. Es kann demnach davon ausgegangen werden, dass besonders in den Klassen, die nach *Chemie im Kontext* unterrichtet wurden, ein deutlicher Schwerpunkt auf fachliche Inhalte gelegt wurde, um dem befürchteten Niveauverlust entgegenzuwirken. Aus diesem Grund schneiden unter Umständen diese Probanden v. a. bei Items mit fachlichem Schwerpunkt gut ab, während sie z. B. gerade den Anforderungen einer Aufgabe, die die Fähigkeit des divergenten Denkens erfordert, nicht so gewachsen ist wie die Klassen ohne *Chemie-im-Kontext*-Erfahrung. In diesem Zusammenhang müssen auch die auffallend guten Leistungen gesehen werden, die von Klassen 9, 11 bis 16 in diesem Test erbracht werden, obwohl zwischen Unterrichtseinheit und der Testbearbeitung die Sommerferien lagen.

Auch wenn es nicht möglich ist, anhand mehrerer Dimensionen Aussagen über einzelne Kompetenzbereiche oder Aufgabenanforderungen zu machen, lassen sich doch Erkenntnisse aus dem Antwortverhalten der Probanden ableiten. In der Analyse konnte keine Gruppe gefunden werden, die sich in allen Aufgabenbereichen durch erfolgreiche Bearbeitung der Aufgaben auszeichnete, vielmehr wurden drei in etwa gleichwertige Gruppen gefunden, die alle einen defizitären Aufgabenbereich zeigen.

Vor dem Hintergrund der Zielformulierung dieser Arbeit muss diskutiert werden, ob und inwieweit die Testitems, die sich als reliabel und aussagekräftig erwiesen haben, in ihrer Kombination den Ansprüchen eines Unterrichts nach *Chemie im Kontext* gerecht werden. Die Aufgaben, die entwickelt und erprobt wurden, sollen sich insbesondere dazu eignen, Kompetenzen, die in einem Unterricht nach *Chemie im Kontext* zentral sind, zur Feststellung des Lernstands von Schülerinnen und Schülern zu erfassen. Es zeigt sich, dass die gemessenen Dimension unterschiedlichste Facetten des naturwissenschaftlichen Arbeitens wiedergibt: Mit den 16 Items der Skala werden alle formulierten Schritte der Erkenntnisgewinnung erfasst, ebenso finden sich alle betrachteten kognitiven Anforderungen in den Aufgaben wieder. Dabei muss angemerkt werden, dass der Bereich der Bewertung im Sinne einer kognitiven Anforderung nur mit einem statt mit

drei Items beschrieben wird, da gerade zwei Aufgaben aus diesem Bereich schlechte Trennschärfen aufweisen und die Reliabilität des Gesamttests verringern.

Dabei muss einschränkend angemerkt werden, dass schriftliche Aufgaben generell nur einen kleinen Ausschnitt des gesamten Spektrums des naturwissenschaftlichen Arbeitens überprüfen können

Abschließend wird diskutiert, wie sich die durch die Testitems gemessene Skala in das Verhältnis von einem Unterricht nach *Chemie im Kontext* und traditionellem Chemieunterricht einordnen lässt. Insgesamt werden durch die Aufgaben zwar alle Prozessschritte des Erkenntnisprozesses erfasst, auf zentrale Aspekte der Unterrichtskonzeption, wie z. B. der Umgang mit Basiskonzepten oder das Formulieren von Fragen, die nur mit Hilfe der Chemie beantwortet werden können, bezieht sich jeweils nur ein Item. Dennoch ist der Fragebogen zur Lernstandserhebung in einem Unterricht nach *Chemie im Kontext* geeignet, da sowohl der Bereich der Planung von Untersuchungen als auch der Bereich der Einordnung der Ergebnisse in den Kontextzusammenhang ein ebenso großes Gewicht erhält wie der Umgang mit ausschließlich fachchemischen Inhalten im Kontext Schulchemie. Bei Betrachtung der Aufgabenanforderungen besitzen Aufgaben, die in erster Linie konvergentes Denken erfordern, erwartungsgemäß einen besonders großen Anteil. Demgegenüber sind Aufgabenanforderungen, die als eher charakteristisch für einen Unterricht nach *Chemie im Kontext* beschrieben werden können (z. B. divergentes Denken und Bewertung), nur in wenigen Aufgaben repräsentiert. Ebenfalls geringen Anteil haben Aufgaben, die den Umgang mit Modellen erfordern, wobei diese Anforderungen in beiden Formen des Chemieunterrichts große Relevanz besitzt. Dadurch dass die gemessene Fragebogendimension Ziele und Anforderungen sowohl eines Unterrichts nach *Chemie im Kontext* als auch eines traditionellen Chemieunterrichts erfasst, lässt sich die Skala nicht eindeutig einem der beiden Unterrichtsformen zuordnen, vielmehr trägt er den Schwerpunkten beider Formen Rechnung.

7 Zusammenfassung und Ausblick

Ziel der Studie war die Entwicklung von kompetenzorientierten Testaufgaben zur Erfassung des Lernstands von Schülerinnen und Schülern am Ende der 10. Klasse, exemplarisch zum Thema „Säuren, Basen und Neutralisation“. Zur Formulierung von Kriterien zur Konstruktion geeigneter Aufgaben wurde ein Modell zum naturwissenschaftlichen Denken und Arbeiten in einem Unterricht nach *Chemie im Kontext* entwickelt und mit kognitiven Anforderungen, die für naturwissenschaftliches Denken und Arbeiten zentral sind, in Beziehung gesetzt. Die im Modell formulierten Kompetenzen und kognitiven Anforderungen sollten im Sinne einer Lernstandserhebung anhand der entwickelten Aufgaben überprüft werden. Dabei ging es einerseits um eine Überprüfung, ob die theoretischen Konstrukte, die der Aufgabenentwicklung zu Grunde gelegt wurden, statistisch bestätigt werden können, und zum anderen um den Gewinn differenzierender Aussagen zum Lernstand von Probanden u. a. mit unterschiedlicher *Chemie-im-Kontext*-Erfahrung. Zudem waren die unterrichtenden Lehrkräfte u. a. dazu aufgefordert, die Leistungen ihrer Lerngruppen bei der Bearbeitung der einzelnen Testaufgaben einzuschätzen.

Im Rahmen der Pilotuntersuchung nahmen 16 Schülerinnen und Schüler an einer ersten, mündlichen Testung der Aufgaben teil, die auf Video aufgezeichnet wurde, darüber hinaus bearbeiteten 120 Schülerinnen und Schüler aus sechs Klassen aus Schleswig-Holstein und Niedersachsen den Fragebogen der Pilotuntersuchung. An der Hauptuntersuchung waren 359 Schülerinnen und Schüler und 17 Lehrkräfte aus Baden-Württemberg, Bayern, Hessen, Niedersachsen und Schleswig-Holstein beteiligt.

Entwicklung von Kategorien zur Analyse und Konstruktion von Aufgaben für einen Unterricht nach Chemie im Kontext

Einleitend wurden die Grundlagen und Zielvorstellungen des Projektes *Chemie im Kontext* vorgestellt. Zur kriteriengeleiteten Entwicklung von Testaufgaben, die insbesondere den Prinzipien und Zielen eines Unterrichts nach *Chemie im Kontext* gerecht wird, wurden im nächsten Schritt Kategorien aufgezeigt, nach denen Aufgaben klassifiziert werden können. Zudem findet sich hier eine Zusammenstellung von Aspekten, die eine Aufgabe als eine *Chemie-im-Kontext*-Aufgabe charakterisieren. Die Entwicklung eines Modells zum naturwissenschaftlichen Denken und Arbeiten in einem Unterricht nach *Chemie im Kontext* und das Zusammenfügen dieser Erkenntnis-schritte mit zentralen kognitiven Anforderungen des naturwissenschaftlichen Arbeitens lieferten Kategorien zur Entwicklung von Aufgaben, die im Rahmen der Arbeit eingesetzt wurden.

Pilotuntersuchung

Aus den Ergebnissen der Pilotstudie ergaben sich wichtige Erkenntnisse zur Überarbeitung der Testitems, des Bewertungsschemas und des Erwartungshorizonts für das Instrument der Hauptstudie. Zu diesem Zweck wurden u. a. Schwierigkeitsindices, Lösungshäufigkeiten, Ankreuzhäufigkeiten einzelner Distraktoren und interne Konsistenz des Messinstruments bestimmt und auf verschiedene Aspekte (z. B. Geschlechtsspezifität, Aufgabenformate) hin analysiert. Die Reliabilitätsanalyse wies das Vorhandensein einer aus zehn Testitems bestehenden Dimension aus, anhand derer sich differenzierende Aussagen zum Abschneiden der Probanden (z. B. Vergleich über alle Klassen) treffen ließen. Insgesamt zeigte die Pilotuntersuchung, dass der Fragebogen als Messinstrument geeignet ist, da er u. a. eine mittlere Schwierigkeit besitzt, in einem zeitlich angemessenen Rahmen bearbeitet werden kann und insgesamt keine geschlechtsspezifische Benachteiligung aufweist.

Hauptuntersuchung

Die Ergebnisse der Hauptstudie zeigten, dass es sich bei dem eingesetzten Schülerfragebogen um ein eindimensionales Testinstrument handelt, mit dem sich die theoretisch zu Grunde gelegten Kategorien nicht getrennt voneinander untersuchen lassen. Vielmehr erfassten die sechzehn Aufgaben eine Skala, die insgesamt das naturwissenschaftliche Denken und Arbeiten wiedergibt. Diese Dimension erstreckt sich von Bereichen des divergenten Denkens bei der Erschließung eines Themenaspektes über die Entwicklung eines Modellexperiments bis hin zur Anwendung von chemischen Konzepten zur Säure-Base-Chemie in verschiedenen Kontexten.

Insgesamt zeigte der unterschiedliche Stand der Vorerfahrung mit der Konzeption *Chemie im Kontext* keinen Effekt auf die Leistungen der teilnehmenden Probanden. Zwischen den teilnehmenden Klassen, die den Fragebogen bearbeiteten, ließen sich dagegen deutliche Unterschiede bei den Testergebnissen feststellen. Zudem konnten für exemplarisch untersuchte Aufgaben häufig Klasseneffekte als Erklärung für das unterschiedliche Abschneiden in den Leistungen ausgemacht werden.

Genaueren Aufschluss über besondere Stärken und Schwächen der teilnehmenden Probanden ergaben einerseits eine Schwierigkeitsanalyse und andererseits eine Clusterzentrenanalyse. Die Clusterzentrenanalyse ermöglichte die Bildung von Probandengruppen, die sich in ihrem Antwortverhalten hinsichtlich verschiedener Gruppen von Aufgaben (z. B. unterrichtstypische Aufgaben, Aufgaben mit eigener Texterstellung, Aufgaben zum Umgang mit graphischen Darstellungen) unterschieden. Hierbei bestätigte sich die Annahme einer Gruppe, die insgesamt dadurch charakterisiert werden kann, dass sie bei allen Aufgabengruppen geringe Lösungserfolge zeigte

(Gruppe der leistungsschwächeren Schüler). Unter den übrigen drei Gruppen fand sich aber keine Probandengruppe, die bei der Bearbeitung aller vier Aufgabengruppen erfolgreich war. Vielmehr zeichneten sich die übrigen drei Gruppen dadurch aus, dass sie jeweils auf einem Gebiet Defizite besaßen. Die Probanden der Gesamtgruppe verteilten sich mit jeweils einem Anteil von etwa 25 % gleichmäßig auf die vier Cluster. Bei Betrachtung der Clusterzugehörigkeit der Probanden nach bestimmten Kriterien (Klassenzugehörigkeit, Geschlecht, *Chemie-im-Kontext*-Erfahrung) ließen sich Unterschiede in der Verteilung ausmachen: In keiner Klasse ließ sich eine ähnlich gleichmäßige Verteilung der Schülerinnen und Schüler über alle vier Cluster wie in der Gesamtstichprobe ausmachen, vielmehr waren einige Klassen in einzelnen Clustern gar nicht oder zu (z. T. deutlich) über 50 % vertreten. Die Analyse der Verteilung von Schülerinnen und Schülern auf die vier Cluster zeigte darüber hinaus, dass die Verteilung insgesamt ein homogeneres Bild zeigte, die Schülerinnen jedoch in den Clustern 2 (Schwierigkeiten mit Aufgaben zur Schulchemie) und 3 (schwache Leistungen in allen Aufgabengruppen) zu einem größeren Prozentsatz vertreten sind, während die männlichen Probanden einen größeren Anteil an den Clustern 1 (schwache Leistungen bei Aufgaben mit modellhafter Erklärung) und 4 (schwache Leistungen bei Aufgaben mit eigener Textproduktion) besaßen. Bei Betrachtung eines Stichprobenauszugs, der sich aufgrund unterschiedlicher *Chemie-im-Kontext*-Erfahrung in zwei Gruppen unterteilen lässt, zeigten sich wiederum stärkere Unterschiede in der Verteilung auf die Cluster. Insbesondere an der Gruppe der eher leistungsschwächeren Schülerinnen und Schüler hatte die Probandengruppe mit *Chemie-im-Kontext*-Erfahrung mit 6 % einen ausgesprochen geringen Anteil, während der Anteil in den Clustern 1 und 4 deutlich über einem Viertel lag, während die Lernenden ohne Konzeptionserfahrung zu 16 % in der Gruppe der leistungsschwachen Schülerinnen und Schüler vertreten waren. Die Anteile dieser Probandengruppe lagen im Cluster 1 etwas unter, in den Clustern 2 und 4 über den Anteilen der Probandengruppe mit Konzeptionserfahrung.

Als weiteres zentrales Ergebnis können die Einschätzungen der Testaufgaben durch die unterrichtenden Lehrkräfte hinsichtlich der Bedeutung der Aufgabenart im Unterricht und der Lösungshäufigkeit in ihrer jeweiligen Klasse gesehen werden. Darüber hinaus erfasste der Lehrerfragebogen, ob die in den Aufgaben erforderten Inhalte und Fähigkeiten Thema des vorangegangenen Unterrichts waren. Es zeigte sich, dass Aufgaben, die als charakteristisch für einen Unterricht, der sich an den Prinzipien von *Chemie im Kontext* orientiert, eingeschätzt wurden, auch für Lehrkräfte, die am Projekt *Chemie im Kontext* teilnahmen, noch nicht in dem Maße von Bedeutung sind wie die eher dem Bereich der traditionellen Schulchemie zugeordneten Aufgaben.

Darüber hinaus legte die Auswertung des Lehrerfragebogens für die einzelnen Klassen z. T. große Differenzen zwischen eingeschätzter und tatsächlicher Lösungshäufigkeit hinsichtlich exemplarisch untersuchter Aufgaben offen. Für die Aufgabe zur Erklärung der Leitfähigkeit, deren zur erfolgreichen Bearbeitung erforderlichen Fähigkeiten und Kenntnisse nach Aussage der Lehrkräfte im Unterricht thematisiert worden waren, lag die von den Lehrkräften geschätzte Lösungshäufigkeit beispielsweise deutlich über der tatsächlichen Lösungshäufigkeit ihrer Schülerinnen und Schüler (z. B. Klasse 1 – von der Lehrkraft geschätzte Lösungshäufigkeit 40 %, tatsächliche Lösungshäufigkeit 25 %; Klasse 11 – von der Lehrkraft geschätzte Lösungshäufigkeit 80 %, tatsächliche Lösungshäufigkeit 40 %). Im Gegensatz dazu unterschätzten die Lehrkräfte die Fähigkeiten ihrer Schülerinnen und Schüler bei einem Aufgabeninhalten (z. B. Bewertung einer Werbung), die nicht im Unterricht bearbeitet wurde (z. B. Klasse 1 – von der Lehrkraft geschätzte Lösungshäufigkeit 45 %, tatsächliche Lösungshäufigkeit 92 %; Klasse 12 – von der Lehrkraft geschätzte Lösungshäufigkeit 20 %, tatsächliche Lösungshäufigkeit 60 %).

Ausblick

Da das entwickelte Testinstrument zwar Unterschiede hinsichtlich verschiedener inhaltlicher Aufgabenanforderungen (z. B. Texterstellung, Umgang mit Graphen) lieferte, die theoretisch intendierten Dimensionen des Modells zum naturwissenschaftlichen Denken und Arbeiten jedoch nicht statistisch getrennt werden konnten, bedarf es diesbezüglich einer weiteren Instrumententwicklung, um Kompetenzbereiche auch zu weiteren chemischen Themen und Kontexten differenziert erfassen zu können. Um Aussagen darüber treffen zu können, in welchen Bereichen eine gezielte Kompetenzförderung notwendig ist, kommt auch der Entwicklung weiterer Instrumente zu anderen Themenbereichen eine große Bedeutung zu. Diese sollten auch dem Anspruch genügen, sowohl die Gütekriterien statistischer Untersuchungen zu erfüllen als auch im Unterricht einsetzbar und gewinnbringend nutzbar zu sein.

Durch die genauere Betrachtung der Diagnosekompetenz der Lehrkräfte wird ferner auf einen weiteren Bereich verwiesen, zu dem ebenfalls weitere Forschungs- und Entwicklungsarbeiten notwendig sind. Ein erster Schritt für die genauere Untersuchung der gemessenen Diskrepanz könnte eine zusätzliche Einschätzung der Lernenden zu der jeweiligen Aufgabenanforderung sein, indem sie angeben müssen, ob sie der Meinung sind, dass die notwendigen Inhalte und Fähigkeiten im Unterricht vermittelt wurden. Ein nachfolgendes Mittel zur Verbesserung der Diagnosekompetenz stellt die Entwicklung von Lehrerfortbildungsmaßnahmen dar, wobei an die im Rahmen von *Chemie im Kontext* bestehenden Konzepte angeknüpft werden kann.

Die im Rahmen dieser Arbeit durchgeführte Untersuchung hatte nicht einen Vergleich von Unterricht nach *Chemie im Kontext* und traditionellem Chemieunterricht zum Ziel. Zu diesem Zweck müsste dem Einsatz des im Rahmen der Arbeit entwickelten Fragebogens eine differenziertere, systematische Erfassung von Unterricht vorangestellt werden, um Unterschiede der beiden Unterrichtsformen sichtbar zu machen und um vergleichende Aussagen treffen zu können. Insbesondere die Lehrereinschätzungen zur Bedeutung einzelner Aufgabenarten im Rahmen der Hauptuntersuchung geben diesbezüglich bereits erste Einblicke, verweisen aber zugleich auf die Schwierigkeiten in der Umsetzung der theoretischen Grundlagen in der Praxis des alltäglichen Chemieunterrichts. Auch hier können ergänzende Entwicklungsarbeiten (z. B. zu weiteren Lehrerhandreichungen und Aufgabenfortbildungen) und Forschungsstudien (z. B. im Rahmen einer Langzeitstudie zur Aufgabeneinbettung) bedeutsame Erkenntnisse für die Weiterentwicklung von Kontextunterricht und Chemieunterricht insgesamt liefern.

LITERATUR

- American Association for the Advancement of Science (AAAS) (1989). Science for All Americans. AAAS Project 2061.
- American Association for the Advancement of Science (AAAS) (1993). Benchmarks for Science Literacy. AAAS Project 2061.
- American Association for the Advancement of Science (AAAS) (2001). Atlas of Science Literacy. AAAS Project 2061.
- Aufschnaiter, C. v. & Aufschnaiter, S. v. (2001). Eine neue Aufgabenkultur für den Physikunterricht. Was fachdidaktische Lernprozess-Forschung zur Entwicklung von Aufgaben beitragen kann. MNU 54/7: 409-416.
- BLK (1997). Expertise "Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts" (http://mail.ipn.uni-kiel.de/projekte/blk_prog/gutacht/index.htm, August 2007).
- BMBF (2007). Zur Entwicklung nationaler Bildungsstandards – Expertise, Bildungsforschung Band 1 (http://www.bmbf.de/pub/zur_entwicklung_nationaler_bildungsstandards.pdf, August 2007).
- Bortz, J. & Döring, N. (2003). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Berlin, Heidelberg, New York: Springer-Verlag.
- Bruder, R. (2003). Konstruieren - auswählen - begleiten. Über den Umgang mit Aufgaben. Friedrich Jahresheft "Aufgaben. Lernen fördern - Selbstständigkeit entwickeln" XXI/2003: 12-15.
- Büchter, A. & Leuders, T. (2005). Mathematikaufgaben selbst entwickeln. Lernen fördern - Leistung überprüfen. Berlin: Cornelsen Verlag Scriptor.
- Bünder, W., Demuth, R. & Parchmann, I. (2003). Basiskonzepte - Welche chemischen Konzepte sollen Schüler kennen und nutzen? Praxis der Naturwissenschaften - Chemie in der Schule 1/52. Jahrgang: 2-7.
- Bünder, W., Demuth, R., Freienberg, J., Klüter, R., Parchmann, I. & Ralle, B. (2006). Lernlinien zur Verknüpfung von Kontextlernen und Kompetenzentwicklung. CHEMKON 13(3): 124-132.
- Bybee, R. W. (2002). Scientific Literacy - Mythos oder Realität? Scientific Literacy. Der Beitrag der Naturwissenschaften zur Allgemeinen Grundbildung. In: W. Gräber, Nentwig, P., Koballa, T., Evans, R. (Hrsg.). Opladen: Leske + Budrich: 21-43.

- Collins, A., Brown, J.S. & Newman, S.E. (1989). *Cognitiv Apprenticeship: Teaching the Crafts of Reading, Writing, and Mathematics. Knowing, Learning and Instruction. Essays in the honor of Robert Glaser.* L. B. H. Resnick. Hillsdale, New Jersey, Erlbaum: 453-494.
- Davier, M. v. & Hansen, H. (1998). "Prüfen: Erfassen und Rückmelden von Kompetenzzuwachs. Erläuterung zu Modul 10 (Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts, SINUS)."
- Deci, E. L. & Ryan, R. M. (1993). "Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik." *Zeitschrift für Pädagogik* 39. Jahrgang, Nr. 2: 223 - 238.
- Demuth, R., Fußangel, K., Gräsel, C., Parchmann, I., Ralle, B., Schellenbach-Zell, J. & Weber, I. (2005). Optimierung von Implementationsstrategien bei innovativen Unterrichtskonzeptionen am Beispiel von "Chemie im Kontext". Schlussbericht.
- Demuth, R., Ralle, B. & Parchmann, I. (2005). Basiskonzepte – eine Herausforderung an den Chemieunterricht. *CHEMKON* 12(2): 55-60.
- Deutsches PISA-Konsortium (2002). Internationales und nationales Rahmenkonzept für die Erfassung von naturwissenschaftlicher Grundbildung in PISA. (<http://www.mpib-berlin.mpg.de/pisa/KurzFrameworkScience.pdf>, August 2007)
- Dörner, D. (1975). *Problemlösen als Informationsverarbeitung.* Stuttgart: Kohlhammer.
- Dubs, R. (1995). Konstruktivismus: Einige Überlegungen aus der Sicht der Unterrichtsgestaltung. *Zeitschrift für Pädagogik* 41. Jahrgang, Nr. 6: 889-903.
- Duit, R. (1995). Zur Rolle der konstruktivistischen Sichtweise in der naturwissenschaftsdidaktischen Lehr- und Lernforschung. *Zeitschrift für Pädagogik*, 41: 905-923.
- Duit, R., Häußler, P. & Prenzel, M. (2001). Schulleistungen im Bereich der naturwissenschaftlichen Bildung. In: *Leistungsmessungen in Schulen.* F. E. H. Weinert. Weinheim, Basel, Beltz Verlag: 169-185.
- Duit, R., Fischer, H. E. & Müller, W. (2002). Vielfalt und Anregung statt Routine. Der Physikunterricht braucht eine andere Aufgabenkultur. *Naturwissenschaften im Unterricht - Physik in der Schule* 13. Jg. (67): 4-7.
- Ellenberger, W. (1993). *Ganzheitlicher Biologieunterricht - Für das Leben lernen.* Berlin: Cornelsen-Verlag.

- Fey, A., Gräsel, C., Puhl, T. & Parchmann, I. (2004). Implementation einer kontextorientierten Unterrichtskonzeption für den Chemieunterricht. *Unterrichtswissenschaft. Zeitschrift für Lernforschung* 33: 238-256.
- Fischer, H. E. & Draxler, D. (2001). Aufgaben und naturwissenschaftlicher Unterricht. *MNU* 54/7: 388-393.
- Fischer, H. E. & Draxler, D. (2002). Konstruktion und Bewertung von Physikaufgaben. In: *Physikdidaktik in der Praxis*, Kircher, E. & Schneider, B., Berlin, Heidelberg, New-York: Springer-Verlag: 300-322.
- Freimann, T. (2003). Bildung? Grundbildung. *Naturwissenschaften im Unterricht – Chemie* 14. Jg., Heft 76/77: 12 - 14.
- Friedrich, F. F. (1997). Selbstgesteuertes Lernen, Lernstrategien und Schule. *Pädagogisches Handeln* 1. Jahrgang, Heft 2: 97-108.
- Gerstenmaier, J. & Mandl, H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik* 41. Jahrgang, Nr. 6: 867-888.
- Graf, D. (2001). Welche Aufgabentypen gibt es? *MNU* 54(7): 422-425.
- Gräsel, C. & Parchmann, I. (2004). Implementationsforschung - oder: der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft. Zeitschrift für Lernforschung* 33: 196-213.
- Hammann, M. (2004). Kompetenzentwicklungsmodelle. Merkmale und ihre Bedeutung - dargestellt anhand von Kompetenzen beim Experimentieren. *MNU* 57/4: 196-203.
- Harp, S. F. & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology* 90 (3): 414-434.
- Hartinger, A., Fölling-Albers, M., Lankes, E.-M., Marenbach, D. & Molfenter, J. (2001). Lernen in authentischen Situationen versus Lernen mit Texten. Zum Aufbau anwendbaren Wissens in der Schriftsprachdidaktik. *Unterrichtswissenschaft. Zeitschrift für Lernforschung* 2: 108-130.
- Häußler, P. & Lind, G. (1998). Weiterentwicklung der Aufgabenkultur im mathematisch-naturwissenschaftlichen Unterricht (Erläuterungen zu Modul 1 mit Beispielen für den Physikunterricht, BLK-Programm "Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts" (1998-2003).
- Häußler, P., Bündler, W., Duit, R., Gräber, W. & Mayer, J. (1998). *Naturwissenschaftsdidaktische Forschung. Perspektiven für die Unterrichtspraxis*. Kiel, IPN.

- Huntemann, H., Vennemann, H., Parchmann, I. (2000). Ein Auto ohne Kunststoffe? - Eine Unterrichtseinheit aus der Konzeption Chemie im Kontext. Praxis der Naturwissenschaften - Chemie in der Schule 4/49. Jg.: 19-30.
- Jäger, M. (2004). Transfer in Schulentwicklungsprojekten. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Klauer, K. J. (1999). Situated Learning: Paradigmenwechsel oder alter Wein in neuen Schläuchen? Zeitschrift für Pädagogische Psychologie 13(3): 117-121.
- Klauer, K. J. (2001). Lehrzielorientierte Tests. Handwörterbuch Pädagogische Psychologie. D. H. H. Rost. Weinheim, Beltz Verlag: 401-405.
- Klieme, E., Artelt, C. & Stanat, P. (2001). Fächerübergreifende Kompetenzen: Konzepte und Indikatoren. Leistungsmessungen in Schulen. F. E. H. Weinert. Weinheim, Basel, Beltz Verlag: 204-218.
- Krapp, A. (1993). Die Psychologie der Lernmotivation. Perspektiven der Forschung und Probleme ihrer pädagogischen Rezeption. Zeitschrift für Pädagogik 39. Jahrgang, Nr. 2: 187-206.
- Kultusministerkonferenz (2004a). Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss (http://www.kmk.org/schul/Bildungsstandards/standards_chemie_30.08.04.pdf, August 2007).
- Kultusministerkonferenz (2004b). Einheitliche Prüfungsanforderungen in der Abiturprüfung Chemie (<http://www.kmk.org/doc/beschl/EPA-Chemie.pdf>, August 2007).
- Lang, M. & Bündler, W. (2004). Developing a Framework for Collaborative Teacher Education Across Subjects. European Development of Innovative Science Teaching (EUDIST), ATEE-Paper.
- Langlet, J. & Freiman, T. (2003). Aufgaben: im Handeln lernen! Unterricht Biologie Heft 287, 27. Jahrgang.
- Lehrke, M. (1998). Verantwortung für das eigene Lernen stärken. Erläuterungen zu Modul 9 (Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts, SINUS).
- Mackensen, I. & Sandmann, A. (2002). Lernen mit Beispielaufgaben: aktiv - konstruktiv - eigenständig. Praxis der Naturwissenschaften - Biologie in der Schule 8/51. Jg: S. 16 - 24.

- Mandl, H., Gruber, H. & Renkl, A. (1995). Situiertes Lernen in multimedialen Lernumgebungen. In: Information und Lernen mit Multimedia. L. J. Issing and P. Klimsa. Weinheim, Psychologie Verlags Union: 167-177.
- Ministerium für Bildung, Wissenschaft, Forschung und Kultur des Landes Schleswig-Holstein (1997). Lehrplan für die Sekundarstufe I der weiterführenden allgemeinbildenden Schulen Hauptschule, Realschule, Gymnasium – Chemie.
- MNU (2000). Chemieunterricht der Zukunft - Qualitätsentwicklung und Qualitätssicherung. Empfehlungen zur Gestaltung von Lehrplänen bzw. Richtlinien für den Chemieunterricht. MNU 53/3: I - XIV.
- Müller, A. & Müller, W. (2002). Physikaufgaben und Kompetenzentwicklung. Naturwissenschaften im Unterricht - Physik in der Schule 13. Jg. (67): 31-33.
- Nentwig, P., Parchmann, I., Demuth, R., Gräsel, C. & Ralle, B. (2002). Chemie im Kontext - From situated learning in relevant contexts to a systematic development of basic chemical concepts. 2nd International Science Education Symposium "Context Based Science Curricula", Strande bei Kiel.
- Nentwig, P., Christiansen, D. & Steinhoff, B. (2004). Eine neue Aufgabenkultur für einen neuen Chemieunterricht. Praxis der Naturwissenschaften - Chemie in der Schule 53. Jahrgang, Nr. 8: 21-24.
- Parchmann, I., Demuth, R., Ralle, B., Paschmann, A. & Huntemann, H. (2001). Chemie im Kontext - Begründung und Realisierung eines Lernens in sinnstiftenden Kontexten. Praxis der Naturwissenschaften - Chemie in der Schule 1/50. Jahrgang: 2-7.
- Parchmann, I., Demuth, R., Ralle, B. & Gräsel, C. (2003). Optimierung von Implementationsstrategien bei innovativen Unterrichtskonzeptionen am Beispiel von "Chemie im Kontext". 1. Zwischenbericht.
- Parchmann, I. & Schmidt, S. (2003). Von erwünschten Verbrennungen und unerwünschten Folgen zum Konzept der Atome. MNU 56 (4): 214-221.
- Parchmann, I. & Menthe, J. (2004). Chemie lernen und mit Chemie lernen - Kompetenzentwicklung im Chemieunterricht. Praxis der Naturwissenschaften - Chemie in der Schule 53 (8): 12-17.
- Pfeifer, P., Lutz, B. & Bader, H. J. (2002). Konkrete Fachdidaktik. Chemie. München, Oldenbourg Schulbuchverlag GmbH.

- Pfeifer, P. (2003). Was heißt "naturwissenschaftliches Arbeiten"? Naturwissenschaften im Unterricht NiU-Chemie 14. Jg., Heft 76/77: 7-11.
- Prechtl, H. & Nerdel, C. (2002). Aufgaben beim computergestützten Lernen dargestellt an einer Simulation zu den Primärreaktionen der Fotosynthese. Praxis der Naturwissenschaften - Biologie in der Schule 8/51: 25-31.
- Prenzel, M., Häußler, P., Rost, J. & Senkbeil, M. (2002). Der PISA-Naturwissenschaftstest: Lassen sich die Aufgabenschwierigkeiten vorhersagen? Unterrichtswissenschaft. Zeitschrift für Lernforschung 30. Jg., Heft 1: 120 - 135.
- Prenzel, M. & Parchmann, I. (2003). Kompetenz entwickeln. Vom naturwissenschaftlichen Arbeiten zum naturwissenschaftlichen Denken. Naturwissenschaften im Unterricht – Chemie 14. Jg., Heft 76/77: 15 - 19.
- Putnam, R. T. & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? Educational Researcher 29 (1): 4-15.
- Renkl, A. (2001). Lernen aus Lösungsbeispielen: Einführung. Unterrichtswissenschaft. Zeitschrift für Lernforschung 1/41.
- Schmidkunz, H., Lindemann, H. (1976 (1. Auflage)). Das Forschend-entwickelnde Unterrichtsverfahren - Problemlösen im naturwissenschaftlichen Unterricht.
- Schmidt, S., Rebentisch, D. & Parchmann, I. (2003). Chemie im Kontext auch für die Sekundarstufe I: Cola und Ketchup im Anfangsunterricht. CHEMKON 10. Jahrgang, Nr. 1: 6-16.
- Simons, P. R. J. (1992). Lernen, selbständig zu lernen - ein Rahmenmodell. Lern- und Denkstrategien. Analyse und Intervention. H. V. f. Psychologie. Göttingen, Mandl, H. & Friedrich, H.F.: 251-264.
- Spiro, R. & Jehng, J.-C. (1990). Cognitive Flexibility and Hypertext: Theory and Technology for the Nonlinear and Multidimensional Traversal of Complex Subject Matter. Nix, D. & Spiro, R.. Cognition, Education, Multimedia – Exploring Ideas in High Technology (163-205). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Stark, R., Gruber, H. & Mandl, H. (1998). Motivationale und kognitive Passungsprobleme beim komplexen situierten Lernen. Psychologie in Erziehung und Unterricht 45. Jg.: 202-215.
- Stäudel, L. (2003a). Naturwissenschaftliches Arbeiten. Naturwissenschaften im Unterricht NiU-Chemie 14. Jg., Heft 76/77: 4 - 6.

- Stäudel, L. (2003b). Der Aufgabencheck. Überprüfen Sie ihre "Aufgabenkultur". Friedrich Jahresheft "Aufgaben. Lernen fördern - Selbstständigkeit entwickeln" XXI/2003: S. 16 - 17.
- Stäudel, L. (2004). Aufgaben nach PISA? - Aufgaben vor PISA! Naturwissenschaften im Unterricht - Chemie in der Schule 15, Nr. 82/83: 91-94.
- Steinhoff, B. (2004). Wissens- und Kompetenzerwerb in einem Unterricht nach Chemie im Kontext – Exemplarische Entwicklung eines Wissenstest für die Unterrichtseinheit (K)Ein Auto ohne Kunststoffe (elektronisch veröffentlichte Dissertation: http://e-diss.uni-kiel.de/diss_1149/ (August 2007)).
- Stracke, I. (2004). Einsatz computerbasierter Concept Maps zur Wissensdiagnose in der Chemie. Empirische Untersuchungen am Beispiel des Chemischen Gleichgewichts. Waxmann Verlag Münster
- Terhart, E. (1999). Konstruktivismus und Unterricht. Gibt es einen neuen Ansatz in der Allgemeinen Didaktik? Zeitschrift für Pädagogik 45. Jahrgang, Nr. 5: 629-647.
- The Cognition and Technology Group at Vanderbilt (1996). Anchored Instruction and Situated Cognition Revisited. In: McLellan, H.. Situated Learning Perspectives (123-154). Englewood Cliffs, New Jersey: Educational Technology Publications.
- Weinert, F. E. H. (2001). Leistungsmessungen in Schulen. Weinheim, Basel, Beltz Verlag.
- Woest, V. (2004). Aufgabenformate. Naturwissenschaften im Unterricht - Chemie in der Schule 15, Nr. 82/83: 7-13.
- Zech, L. K., Gause-Vega, C. L., Bray, M. H., Secules, T. & Goldman, S. R. (2000). "Content-based collaborative inquiry: A professional development model for sustaining educational reform." Educational Psychologist 35(207-217).

Anhang

Die Fragebögen sind dem Format der Arbeit angepasst und in ihrer Darstellung auf das Wesentliche beschränkt.

A Schülerfragebogen der Pilotstudie

Liebe Schülerin, lieber Schüler!

In diesem Test findest du eine Reihe von Aufgaben und Fragen zu unterschiedlichen Themen. Einige Aufgaben sind kurz und andere sind etwas länger, ein paar Aufgaben werden dir schwerer fallen und andere werden leicht sein, einige haben viel, andere auf den ersten Blick weniger mit Chemie zu tun. Bitte versuche, alle Aufgaben so gut wie möglich zu lösen.

Dir werden verschiedene Typen von Aufgaben begegnen:

Bei einigen Aufgaben musst du deine Antwort zeichnen oder als kurzen Text oder Rechnung aufschreiben. Wenn der Platz auf dem Zettel nicht ausreicht, nimm bitte die Rückseite!

Bei anderen Aufgaben kreise bitte die richtigen Lösungen ein.

Einige Fragen zu deiner Person:

Wie alt bist du? _____ Jahre

Geschlecht? männlich weiblich

Verrätst du uns deine letzte Chemienote? Im letzten (Halbjahres-) Zeugnis war es eine

- 1
- 2
- 3
- 4
- 5
- 6

Hast du vor, in der Oberstufe Chemie zu belegen?

- nein
- ja, als Grundkurs
- ja, als Leistungskurs
- weiß ich noch nicht

Vielen Dank für deine Mithilfe!

Sauer macht lustig?! Säuren machen Spaß?!

Nenne vier Stichworte, die dir aus deinem Chemieunterricht zum Thema „Säuren“ einfallen!

a) _____ b) _____

c) _____ d) _____

I. Was tun, wenn's im Magen brennt?



Dein Magen enthält eine saure Lösung, die die Verdauung unterstützt und Bakterien, die mit dem Essen in den Magen gelangen, abtötet. Wenn im Magen zu viel Säure produziert wird, dann „brennt“ es – z.B. in der Speiseröhre, wenn man Aufstoßen muss. Dies wird als „Sodbrennen“ bezeichnet.

I.1. Wenn du eine Tablette gegen die Übersäuerung des Magens herstellen solltest, **bei welchen Stoffen würdest du nach einer geeigneten Substanz suchen?**

- A Säuren
- B Laugen
- C sauer wirkende Salze
- D basisch wirkende Salze
- E Aktivkohle

I.2. Ganz unterschiedliche Aspekte müssen bei der Entwicklung eines neuen Medikaments berücksichtigt werden. **Nenne fünf Kriterien, die dir bei der Entwicklung eines Medikaments wichtig erscheinen!**

a) _____ b) _____ c) _____

d) _____ e) _____

I.3. Du möchtest dich über ein vorliegendes Medikament gegen Sodbrennen genauer informieren. Dazu hast du Gelegenheit, die folgenden Personen zu befragen.

Formuliere je eine Frage zu dem Medikament an

- deine Chemielehrerin/deinen Chemielehrer:
-

- deine Biologielehrerin/deinen Biologielehrer:
-

- eine Apothekerin/einen Apotheker:
-

Eine große Fülle von Informationen liefert auch das Internet.

Mit welchen drei Suchbegriffen startest du hier deine Recherche?

a) _____ b) _____ d) _____

I.4. Nach der Einnahme von Bullrich-Salz muss man häufig aufstoßen. Grund ist die folgende chemische Reaktion im Magen: Der Wirkstoff Natriumhydrogencarbonat bildet mit der im Magensaft enthaltenen Salzsäure Kohlenstoffdioxid.



Bullrich-Salz® ist der Klassiker unter den „Magensäurestoppersn“!

Anwendung: Bei Sodbrennen, Magendruck, Völlegefühl.

Inhaltsstoffe: Natriumhydrogencarbonat, Talkum, Maisstärke

Überlege dir einen Modellversuch, mit dem du verdeutlichen kannst, was mit einer Bullrich-Salz-Tablette im Magen passiert und warum man nach ihrer Einnahme aufstoßen muss. Das entstehende Gas soll aufgefangen werden, um damit weitere Versuche durchführen zu können.

Erstelle eine Zeichnung der Versuchsanordnung und beschrifte sie ausführlich!

(Platz zum Zeichnen hast du auf der folgenden Seite!)

Versuchszeichnung:

- I.5. Der Packungsbeilage von Bullrich-Salz ist zu entnehmen, dass in einer Tablette 840 mg Natriumhydrogencarbonat enthalten sind.

Formuliere die Reaktionsgleichung für die Neutralisationsreaktion von NaHCO_3 mit HCl !

Gib an, wieviel Mol 840mg NaHCO_3 (enthalten in einer Tablette) entsprechen!

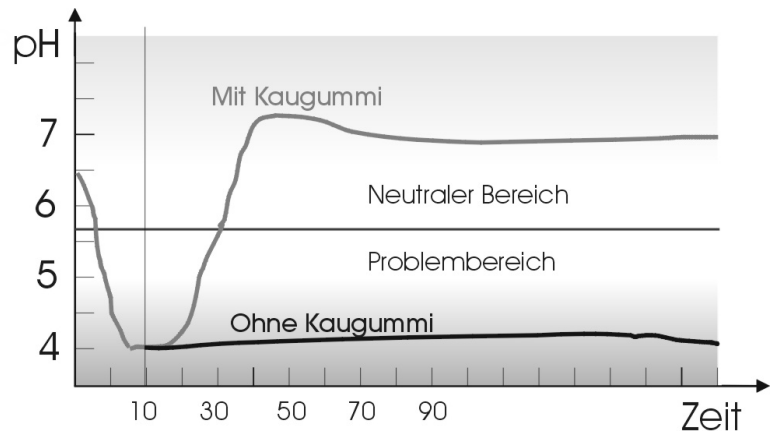
Berechne nun, wieviel HCl (in g) eine Bullrich-Salz-Tablette neutralisieren kann!

Hinweis: Bei der Berechnung der molaren Massen hilft dir das abgebildete Periodensystem!

1 H 1,0079							2 He 4,0026
3 Li 6,94	4 Be 9,0122						
11 Na 22,990	12 Mg 24,305	5 B 10,811	6 C 12,011	7 N 14,007	8 O 15,999	9 F 18,998	10 Ne 20,180
19 K 39,098	20 Ca 40,078	13 Al 26,982	14 Si 28,086	15 P 30,974	16 S 32,066	17 Cl 35,453	18 Ar 39,948
37 Rh	38 Sr	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr

II. Kaugummi gegen Karies

Nach der Nahrungsaufnahme entstehen durch Bakterien im Mund Säuren. Diese Säuren schädigen den Zahnschmelz, die Zähne werden anfällig für Karies. Eine Kaugummi-Werbung verspricht, durch ein speziell entwickeltes Zahnpflege-Kaugummi das Problem der Kariesanfälligkeit der Zähne nach dem Essen zu lösen. Mit der nebenstehenden Grafik wirbt das Unternehmen für seine Kaugummis. Der Liste der Inhaltsstoffe auf der Kaugummi-Verpackung ist zu entnehmen, dass in diesen Kaugummis kein basisch reagierender Stoff enthalten ist.



II.1. Sieh dir die abgebildete Grafik an! Kennzeichne in dem Kurvenverlauf zunächst den Punkt, an dem das Kauen des Kaugummis zu wirken beginnt!

Beschreibe den Verlauf der Kurve: was passiert nach Aussage des Kaugummi-Herstellers, wenn du nach dem Essen eines seiner Kaugummis kaust!

II.2. Sieh dir die Abbildung noch einmal an: Solche Bilder sollen vor allem werbewirksam sein. Manchmal lohnt es sich aber, ein bißchen genauer hinzuschauen.

Worüber stolpert dein „chemischer Sachverstand“ in dieser Grafik?

II.3. Geht deinen Eltern „Kaugummikauen“ auch auf die Nerven?

Mit welchen der folgenden Argumente kannst du erklären, dass du deinen Zähnen mit dem „Kaugummikauen“ etwas Gutes tust? Kennzeichne „ja“ oder „nein“!

Ist das ein Argument zu Gunsten des „Kaugummikauens“?	Ja oder nein?
Zucker im Kaugummi konserviert die Zähne.	ja / nein
Durch Kauen angeregter Speichelfluss reinigt die Zähne.	ja / nein
Das Kauen eines Kaugummis reinigt die Zähne.	ja / nein
Häufiges Kauen trainiert die Zähne.	ja / nein

II.4. In einer Studie wurden 225 Personen gebeten, drei Monate lang täglich drei Mal ein Zahnpflege-Kaugummi zu kauen. Die Testpersonen wurden vor und nach dieser Zeit zahnärztlich untersucht. Bei 175 Personen war der Zustand unverändert, bei 40 hatte er sich leicht verbessert, und bei den restlichen 10 geringfügig verschlechtert.

Welche der folgenden Aussagen ist korrekt?

- A Diese Studie beweist, dass Kaugummikauen gut für die Zähne ist.
- B Diese Studie beweist, dass Kaugummikauen keine Wirkung auf die Zähne hat.
- C Diese Studie beweist, dass Kaugummikauen schlecht für die Zähne ist.
- D Diese Studie beweist gar nichts.



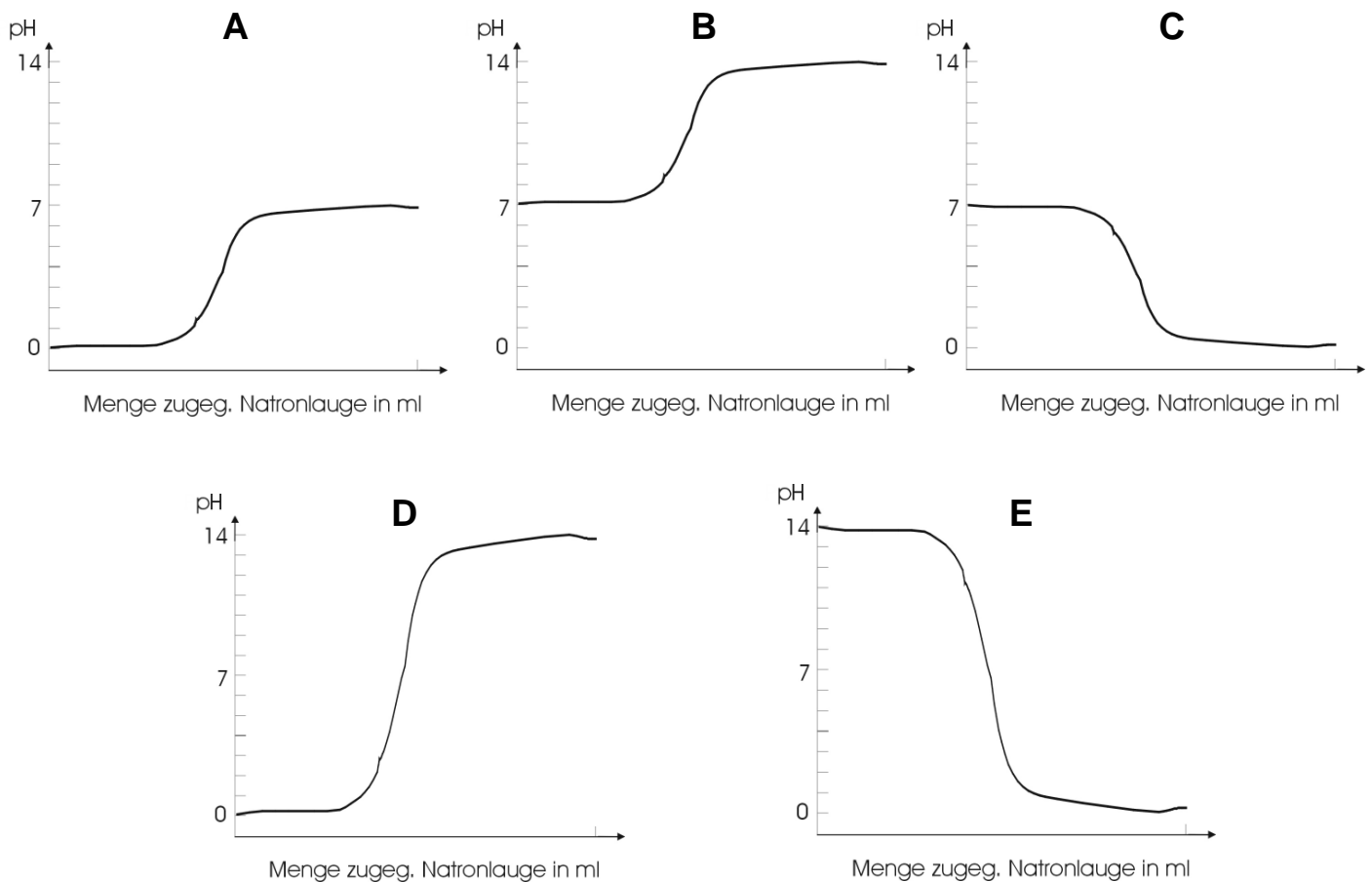
Bitte Umblättern - jetzt kommt der letzte Teil des Tests!

III. Im Chemielabor

III.1. Der Chemikalienschrank ist gut sortiert. In welcher Aufzählung werden ausschließlich Stoffe genannt, deren wässrige Lösungen basisch (alkalisch) reagieren?

- A NaOH, CH₃COOH, H₂SO₄
- B NaOH, CaO, NaHCO₃
- C NaCl, NH₃, KOH
- D CH₃CH₂OH, CaO, HCl

III.2. In einem Gefäß befindet sich Salzsäure. Zu der Salzsäure wird tropfenweise Natronlauge zugegeben. Diese Natronlauge besitzt die gleiche Konzentration wie die Salzsäure, ist aber in der doppelten Menge vorhanden. Nach und nach wird die gesamte Natronlauge zugetropft: **Kennzeichne, welche der folgenden Kurven die pH-Wert-Änderung richtig wiedergibt!**



- III.3.** Beim Experimentieren wurde im Abzug Säure verschüttet. Eine Mitschülerin schlägt vor, Wasser hinzuzugeben, ein anderer Mitschüler möchte Natronlauge nehmen.
Welchem Ratschlag schließt du dich an? Begründe deine Entscheidung!

- III.4.** Indikatoren verändern je nach pH-Wert einer Lösung ihre Farbe. **Kennzeichne, mit welchem der drei Indikatoren du prüfen kannst, ob Regenwasser pH-neutral ist!**

A	Bromthymolblau	Umschlagbereich bei pH 6,2 (gelb) – 7,6 (blau)
B	Phenolphthalein	Umschlagbereich bei pH 8,0 (farblos) – 9,8 (violett)
C	Methylorange	Umschlagbereich bei pH 3,1 (rot) – 4,4 (orange)

- III.5.** **Trage in die linke Spalte der Tabelle** zwei Eigenschaften ein, die allen sauren Lösungen gemeinsam sind!
Trage dann in der rechten Spalte der Tabelle je einen Versuch ein, mit dem du die genannte Eigenschaft demonstrieren kannst!

Eigenschaft / Merkmal	geeigneter Versuch zur Demonstration
1. Eigenschaft einer sauren Lösung:	
2. Eigenschaft einer sauren Lösung:	

- III.6.** Im Unterricht hast du Modelle kennen gelernt, mit denen du das typische Verhalten von Säuren erklärt kannst. Ein Beispiel:
Reines Wasser leitet den elektrischen Strom praktisch nicht.
Salzsäure hingegen leitet den elektrischen Strom gut.

Erläutere bitte das Phänomen!

Vielen Dank für deine Mitarbeit und dein Durchhaltevermögen!



B Erwartungshorizont zur Bewertung der Aufgaben mit offenem Antwortformat (Pilotstudie)

Aufgabe	Erwartungshorizont (für mehrstufiger Auswertung wird eine Stufung Antworten angegeben)				
I.2	Für jedes der fünf angegebenen Kriterien wird entscheiden, ob es sinnvoll ist; „sinnvoll“ ist im Aufgabenzusammenhang: Preis, Wirtschaftlichkeit, einfache, kostengünstige Herstellung, keine Giftwirkung, Nebenwirkungen, Verträglichkeit, Wirkung, Geschmack, einfache Aufnahme/Einnahme, dabei werden „Verträglichkeit“, „schädliche Wirkung auf den Körper“ und „Nebenwirkungen“ als drei richtige Kriterien gewertet.				
I.3	Folgende Fragen werden als richtig bewertet: <i>Chemielehrerin/Chemielehrer:</i> z.B. Welche Stoffe sind in dem Medikament enthalten? Wie wirkt das Medikament? Welche chemischen Eigenschaften hat der Wirkstoff? Welche chemische Struktur hat der Wirkstoff/das Medikament? Welche Summenformel hat der Wirkstoff? Welche chemischen Reaktionen laufen im Körper ab? <i>Biologielehrerin/Biologielehrer:</i> z.B. Warum ist der pH-Wert im Magen so niedrig? Wozu brauchen wir die Säure im Magen? Was passiert mit dem Medikament im Magen/Körper? Wie wird das Medikament abgebaut? Wie wirkt das Medikament? (Fragen zum Abbau im Körper, Weg im Körper, biologischen Ursprung des Medikaments) <i>Apothekerin/Apotheker:</i> z.B. Wie wirkt das Medikament? Welche Nebenwirkungen hat das Medikament? Würden Sie das Medikament empfehlen? Welche Medikamente gegen Sodbrennen gibt es? Was kosten die verschiedenen Medikamente? Was empfehlen Sie mir?				
I.3b	Es werden alle Suchbegriffe als richtig gewertet, wenn sie einen Bezug zur Krankheit oder zum Medikament haben, z.B. Sodbrennen, Magensäure, Medikamente, Bullrich Salz, Neutralisation, Aufstoßen, Nebenwirkungen, Hersteller, Name des Medikaments, ...				
I.4	Bewertung der Versuchszeichnung: 2 Punkte: Die Zeichnung gibt einen sinnvollen Versuchsaufbau wieder. Aus der Zeichnung wird deutlich, wie das Gas aufgefangen (z.B. Luftballon) und abgeleitet wird und worin sich das Reaktionsgemisch befindet. Das Kohlenstoffdioxid muss nicht rein aufgefangen werden, für Nachweise reicht ein Gemisch mit Luft. Darüber hinaus wird unterstellt, dass die Lernenden wissen, dass das aufgefangene Gas schwerer als Luft ist: Ein nach oben offenes Gefäß funktioniert als Auffanggefäß für das entstehende Gas. Das Gas wird so aufgefangen, dass damit weitere Versuche durchgeführt werden können (Glimmspan, Kerze, Barytwasser, es geht nicht um eine quantitative Bestimmung). Die eingesetzten Gefäße/Geräte und die Chemikalien sind beschriftet. 1 Punkt: wie unter Punkt 2 beschrieben mit folgenden Schwächen. Das Gas wird direkt in Kalkwasser geleitet und nicht aufgefangen oder die Zeichnung ist unsauber, verdeutlicht aber den Gedanken des Auffangens richtig. 0 Punkte: Die Zeichnung gibt ein geschlossenes System wieder oder ist grundsätzlich fehlerhaft.				
I.5.	Die Reaktionsgleichung muss vollständig korrekt sein. Für die Berechnung wird festgelegt: 2 Punkte: Die Berechnung wird vollständig und korrekt durchgeführt. 1 Punkt: Die Molmasse von NaHCO_3 wird bestimmt und korrekt umgerechnet (840 mg).				
II.1.	Falsche Markierungen in der Kurve werden nicht bewertet, können aber Aufschluss über das Verständnis der Kurve geben. Die Antwort sollte folgende inhaltlichen Aspekte enthalten: Nach dem Essen liegt der pH-Wert im Mund bei 4 (saurer Bereich). Durch das Kauen eines Kaugummis steigt der pH-Wert laut Grafik wieder an bis in den ungefährlichen pH-Bereich, die Kurve steigt bis pH 7 (pH-neutral) und bleibt dann konstant. Wichtig für eine positive Bewertung ist, dass der Proband einen expliziten Bezug zur Grafik herstellt, z.B. durch Verweis auf die Achsen oder Beschreibung des Kurvenverlaufs.				
II.2.	Für eine positive Bewertung wird die Nennung eines Aspektes, z.B. unterschiedliche Begriffsdefinition von pH-neutral in Chemie und Werbung; Zweifel an der dargestellten Information, dass der pH-Wert im Mund über einen recht langen Zeitraum bei ca. 4 liegt; Lage des pH-Werts vor dem Essen bei 6,5, nach dem Kauen des Kaugummis bei ca. 7, fehlende Zeitangabe.				
III.3	2 Punkte: Entscheidet sich nicht für Natronlauge, da man zuvor genau die Menge Natronlauge berechnen müsste, die man benötigt, um die Säure zu neutralisieren. Entscheidet sich für (viel) Wasser, um die Säure (stark) zu verdünnen (und damit pH-Wert zu erhöhen). Entscheidet sich für Natronlauge, verweist aber darauf, dass der pH-Wert kontrolliert werden muss, damit man nicht zuviel Natronlauge zugibt. [Ein Verweis auf die freiwerdende Neutralisationswärme als zusätzliche Gefahrenquelle ist nicht notwendig zur vollständigen und korrekten Lösung dieser Aufgabe.] 1 Punkt: Entscheidet sich für Natronlauge, weil es die Säure neutralisiert. 0 Punkte: Schließt sich keinem Ratschlag an oder begründet Entscheidung nicht.				
III.5.	<table border="1" style="width: 100%;"> <thead> <tr> <th>Eigenschaft/Merkmal</th> <th>geeigneter Versuch zur Demonstration</th> </tr> </thead> <tbody> <tr> <td>pH-Wert < 7; färbt Indikatoren charakteristisch; Reaktion mit Metallen, Reaktionen mit Kalk; ätzende Wirkung allgemein</td> <td>Nachweis mit pH-Papier/ Indikatorlösungen → Verfärbung; Gasbildung mit unedlen Metallen/Kalk</td> </tr> </tbody> </table> 2 Punkte: Nennt zwei richtige Eigenschaften, davon eine mit geeignetem Versuch. 1 Punkt: Nennt eine richtige Eigenschaft mit einem geeigneten Versuch oder zwei richtige Eigenschaften.	Eigenschaft/Merkmal	geeigneter Versuch zur Demonstration	pH-Wert < 7; färbt Indikatoren charakteristisch; Reaktion mit Metallen, Reaktionen mit Kalk; ätzende Wirkung allgemein	Nachweis mit pH-Papier/ Indikatorlösungen → Verfärbung; Gasbildung mit unedlen Metallen/Kalk
Eigenschaft/Merkmal	geeigneter Versuch zur Demonstration				
pH-Wert < 7; färbt Indikatoren charakteristisch; Reaktion mit Metallen, Reaktionen mit Kalk; ätzende Wirkung allgemein	Nachweis mit pH-Papier/ Indikatorlösungen → Verfärbung; Gasbildung mit unedlen Metallen/Kalk				
III.6	2 Punkte: Nennt die Existenz von Ionen in der Lösung als Voraussetzung für elektrische Leitfähigkeit und erklärt die Leitfähigkeit von sauren Lösungen über Bildung von H_3O^+ -Ionen/über die Bildung von Protonen und Säure-Anionen (Säuren sind Protonendonatoren). 1 Punkt: Die Begründung ist nicht vollständig, aber korrekt: z.B. „durch vorliegende Ionen kann Strom geleitet werden“ oder „Säure bildet in Wasser Ionen“.				

Tab. 38: Erwartungshorizont zur Bewertung der Aufgaben mit offenem Antwortformat der Pilotstudie, bei mehrstufigen Items ist für jede Stufung die erforderliche Bearbeitung beschrieben

C Klassifizierung der Testaufgaben (Pilotstudie)

Ausprägung Aufgabe	Oberflächenstruktur			Wissensbestände					Antwortformat						
	Sodbrennen	Kaugummi	Schulchemie	pH-Wert, Indikator	Säuren (Eig.)	Basen	Neutralisation	keine der angeg.	Begriffe, Stichwort	Text	Zeichnung	Rechnung	Reaktionsgl.	MC (einfach)	MC (kompl.)
„Gegenmittel Sodbrennen“	X						X							X	
„Aspekte der Medikamententwicklung“	X							X	X						
„Fragen formulieren“	X							X		X					
„Internet-Suchbegriffe“	X							X	X						
„Versuch zum Gasauffangen“	X							X			X				
„Reaktionsgleichung“	X						X						X		
„Berechnung“	X							X				X			
„Grafik beschreiben“		X		X						X					
„Fehlersuche“		X		X						X					
„Kaugummivirkung“		X						X							X
„Studie bewerten“		X						X						X	
„basische Stoffe“			X			X								X	
„pH-Kurven“			X	X										X	
„Ratschlag“			X				X			X					
„Indikatoren“			X	X										X	
„Demonstrationsversuche“			X		X					X					
„Leitfähigkeit“			X		X					X					

Tab. 39a: Kategorisierung der Testaufgaben der Pilotuntersuchung anhand der ausgearbeiteten Kriterien, Teil 1.

Anhang

Ausprägung Aufgabe	Erkenntnissschritte								kognitive Aufgabenanforderung						
	Kontext erschließen	chem. Fragen	Planung	Durchfüh- rung	Ergebnisse	Einordnung (Kontext)	Einordnung (Chemie)	Basiskon- zepte	div. Denken	konv. Denken	Zahlen	Graphen	Sprache	Modelle	Bewertung
„Gegenmittel Sodbrennen“						X				X					
„Aspekte der Medikamen- tentwicklung“	X								X						
„Fragen formulieren“		X											X		
„Internet- Suchbegriffe“			X							X					
„Versuch zum Gas auffangen“			X							X					
„Reaktions- gleichung“					X								X		
„Berechnung“				X							X				
„Grafik beschreiben“					X							X			
„Fehlersuche“						X									X
„Kaugummi- wirkung“						X									X
„Studie bewerten“						X									X
„basische Stoffe“							X			X					
„pH-Kurven“							X					X			
„Ratschlag“							X			X					
„Indikatoren“							X			X					
„Demonstra- tionsversuche“					X					X					
„Leitfähigkeit“								X						X	

Tab. 39b: Kategorisierung der Testaufgaben der Pilotuntersuchung anhand der ausgearbeiteten Kriterien, Teil 2.

D Pilotstudie: Ergänzungen zum ErgebnisteilFragestellung 2.1 (Fragebogeneinsatz)

Testaufgabe	N	M	SD	Trennschärfe
I.1*	117	1,32	0,597	0,39
I.2*	112	1,24	0,852	0,41
I.3a*	120	1,58	0,616	0,07
I.3b*	118	1,74	0,530	0,15
I.4*	91	0,52	0,689	0,10
I.5a	77	0,43	0,498	0,50
I.5b*	21	0,29	0,717	- 0,04
II.1	117	0,25	0,434	0,10
II.2	92	0,49	0,503	0,40
II.3	120	0,24	0,430	0,04
II.4	120	0,48	0,502	0,13
III.1	101	0,43	0,497	0,28
III.2	112	0,67	0,472	0,40
III.3*	112	0,92	0,699	0,39
III.4	108	0,71	0,454	0,45
III.5*	91	0,89	0,823	0,56
III.6*	79	0,57	0,614	0,48

Tab. 40: Kenndaten zur internen Konsistenz des **Gesamttests**; **Cronbachs $\alpha = 0,70$** ; für die Bestimmung der Trennschärfen sind fehlende Antworten als falsch gewertet worden, daher $N = 120$; für die in der Tabelle angegebenen Kenndaten werden die Werte für die tatsächlich bearbeiteten Aufgaben (N) wiedergegeben; die mehrstufig ausgewerteten Aufgaben sind durch Sternchen gekennzeichnet.

Testaufgabe (10 Items)	N	M	SD	Trennschärfe
I.1*	117	1,32	0,597	0,42
I.2*	112	1,24	0,852	0,45
I.5a	77	0,43	0,498	0,59
II.2	92	0,49	0,503	0,45
III.1	101	0,43	0,497	0,29
III.2	112	0,67	0,472	0,40
III.3*	108	0,71	0,454	0,35
III.4	112	0,92	0,699	0,47
III.5*	91	0,89	0,823	0,57
III.6*	79	0,57	0,614	0,53

Tab. 41: Kenndaten zur internen Konsistenz des Tests, wenn nur die zehn Items berücksichtigt werden, deren Trennschärfe $> .20$ ist; **Cronbachs $\alpha = 0,77$** ; für die Bestimmung der Trennschärfen sind fehlende Antworten als falsch gewertet worden, daher $N = 120$; für die in der Tabelle angegebenen Kenndaten werden die Werte für die tatsächlich bearbeiteten Aufgaben (N) wiedergegeben; die mehrstufig ausgewerteten Aufgaben sind durch Sternchen gekennzeichnet

Fragestellung 2.2 (Fragebogeneinsatz)

Testaufgabe (MC-Format)	Distraktor	N	Ankreuzhäufigkeit	Lösungshäufigkeit gesamt
I.1*	A	117	1,7	38,5 (54,7)
	B	117	(57,3)	
	C	117	4,3	
	D	117	57,3	
	E	117	1,7	
		Σ 117		
II.3 (Zustimmung zur Aussage)	A	116	1,7	24,2
	B	115	58,3	
	C	113	59,3	
	D	115	20,0	
		Σ 120		
II.4	A	120	6,7	48,3
	B	120	45,0	
	C	120	1,7	
	D	118	54,2	
		Σ 120		
III.1	A	101	21,8	42,6
	B	101	46,5	
	C	101	27,7	
	D	101	12,9	
		Σ 101		
III.2	A	112	16,1	67,0
	B	112	5,4	
	C	112	8,0	
	D	112	67,0	
	E	112	4,5	
		Σ 112		
III.4	A	108	71,3	71,3
	B	108	19,4	
	C	108	9,3	
		Σ 108		

Tab. 42: Verteilung der Ankreuzhäufigkeiten der einzelnen Distraktoren aller sechs Multiple-Choice-Aufgaben; angegeben sind die Anzahl der tatsächlich bearbeiteten Aufgaben (N) und die Lösungshäufigkeit insgesamt, die Lösungshäufigkeit der teilrichtigen Lösungen ist in Klammern angegeben.

Fragestellung 2.3 (Fragebogeneinsatz)

Testaufgabe	Geschlechtzugehörigkeit	N	M	SD	df	t	p
I.1 *	männlich	72	1,39	0,618	101,89	1,726	n.s.
	weiblich	45	1,20	0,548			
I.2 *	männlich	69	1,26	0,816	110	0,310	n.s.
	weiblich	43	1,21	0,914			
I.3a *	männlich	73	1,47	0,689	117,89	- 2,94	< 0,01 ²
	weiblich	47	1,77	0,428			
I.3b *	männlich	72	1,69	0,597	116	- 1,10	n.s.
	weiblich	46	1,80	0,401			
I.4 *	männlich	56	0,45	0,658	89	- 1,23	n.s.
	weiblich	35	0,63	0,731			
I.5a	männlich	45	0,49	0,506	75	1,27	n.s.
	weiblich	32	0,34	0,483			
I.5b *	männlich	12	0,33	0,778	19	0,34	n.s.
	weiblich	9	0,22	0,667			
II.1	männlich	72	0,24	0,428	115	- 0,37	n.s.
	weiblich	45	0,27	0,447			
II.2	männlich	60	0,50	0,504	90	0,28	n.s.
	weiblich	32	0,47	0,507			
II.3	männlich	73	0,21	0,407	118	- 1,15	n.s.
	weiblich	47	0,30	0,462			
II.4	männlich	73	0,53	0,502	118	1,39	n.s.
	weiblich	47	0,40	0,496			
III.1	männlich	60	0,47	0,503	99	1,00	n.s.
	weiblich	41	0,37	0,488			
III.2	männlich	69	0,70	0,464	110	0,74	n.s.
	weiblich	43	0,63	0,489			
III.3 *	männlich	71	0,97	0,696	110	1,04	n.s.
	weiblich	41	0,83	0,704			
III.4	männlich	68	0,74	0,444	106	0,66	n.s.
	weiblich	40	0,68	0,474			
III.5 *	männlich	52	0,87	0,841	89	- 0,33	n.s.
	weiblich	39	0,92	0,807			
III.6 *	männlich	52	0,48	0,577	77	- 1,81	n.s.
	weiblich	27	0,74	0,656			

Tab. 43: Unterschiede im Bearbeitungserfolg bei männlichen und weiblichen Probanden bei einem Signifikanzniveau $p \leq 0,01$; Mittelwert des prozentualen Anteils richtig gelöster Aufgaben, Standardabweichung, Minimum und Maximum; ²Effektstärke $d = 0,52$; mehrstufige Items sind mit einem Sternchen gekennzeichnet.

Fragestellung 2.3 (Fragebogeneinsatz)

Testaufgabe / Distraktoren	Geschlechtzugehörigkeit	N	M	SD	df	t	p	
I.1	A	männlich	72	0,01	0,118	115	- 0,336	n.s.
		weiblich	45	0,02	0,149			
	B	männlich	72	0,51	0,503	97,37	- 1,650	n.s.
		weiblich	45	0,67	0,477			
	C	männlich	72	0,04	0,201	115	- 0,072	n.s.
		weiblich	45	0,04	0,208			
	D	männlich	72	0,63	0,488	115	1,449	n.s.
		weiblich	45	0,49	0,506			
	E	männlich	72	0,03	0,165	115	1,124	n.s.
		weiblich	45	0,00	0,000			
II.3	A	männlich	69	0,01	0,120	114	- 0,273	n.s.
		weiblich	47	0,02	0,146			
	B	männlich	71	0,55	0,501	113	- 0,916	n.s.
		weiblich	44	0,64	0,487			
	C	männlich	69	0,52	0,503	97,48	- 1,982	n.s.
		weiblich	44	0,70	0,462			
	D	männlich	70	0,21	0,413	113	0,474	n.s.
		weiblich	45	0,18	0,387			
II.4	A	männlich	73	0,05	0,229	118	- 0,645	n.s.
		weiblich	47	0,09	0,282			
	B	männlich	73	0,41	0,495	118	- 1,068	n.s.
		weiblich	47	0,51	0,505			
	C	männlich	73	0,03	0,164	118	1,141	n.s.
		weiblich	47	0,00	0,000			
	D	männlich	73	0,56	0,500	116	0,531	n.s.
		weiblich	45	0,51	0,506			
III.1	A	männlich	60	0,22	0,415	99	- 0,34	n.s.
		weiblich	41	0,22	0,419			
	B	männlich	60	0,53	0,503	99	1,663	n.s.
		weiblich	41	0,37	0,488			
	C	männlich	60	0,32	0,469	99	1,067	n.s.
		weiblich	41	0,22	0,419			
	D	männlich	60	0,08	0,279	65,87	- 1,547	n.s.
		weiblich	41	0,20	0,401			
III.2	A	männlich	69	0,17	0,382	110	0,478	n.s.
		weiblich	43	0,14	0,351			
	B	männlich	69	0,04	0,205	110	- 0,596	n.s.
		weiblich	43	0,07	0,258			
	C	männlich	69	0,06	0,235	110	- 1,100	n.s.
		weiblich	43	0,12	0,324			
	D	männlich	69	0,70	0,464	110	0,736	n.s.
		weiblich	43	0,63	0,489			
	E	männlich	69	0,04	0,205	110	- 0,075	n.s.
		weiblich	43	0,05	0,213			
III.4	A	männlich	68	0,74	0,444	106	0,664	n.s.
		weiblich	40	0,68	0,474			
	B	männlich	68	0,18	0,384	106	- 0,611	n.s.
		weiblich	40	0,23	0,423			
	C	männlich	68	0,09	0,286	106	- 0,202	n.s.
		weiblich	40	0,10	0,304			

Tab. 44: Unterschiede im Ankrenzverhalten bei Distraktoren der Multiple-Choice-Aufgabe bei männlichen und weiblichen Probanden bei einem Signifikanzniveau $p \leq 0,01$: Mittelwert des prozentualen Anteils richtig gelöster Aufgaben, Standardabweichung, Minimum und Maximum.

Fragestellung 2.5 (Fragebogeneinsatz)

Klasse	N	M	SD	Min.	Max.
1	25	7,76	1,71	4,5	11,5
2	11	6,18	1,19	4,5	8,00
3	18	11,64	1,86	8,0	14,5
4	18	6,83	1,87	3,5	11,0
5	22	6,43	2,36	1,5	11,0
6	26	5,31	2,33	1,0	11,0
Σ 120		7,28	2,81	1,0	14,5

Tab. 45: Abschneiden der einzelnen Klassen 1 – 6 im Gesamttest (bei einer maximal erreichbaren Punktzahl von 17): Klassengröße N, Mittelwert der insgesamt erreichten Punktzahlen im Test, Standardabweichung, Minimum und Maximum

Diskussion zur Fragestellung 2.2 (Fragebogeneinsatz)

Klasse	N	fehlend	Zustimmung in gültige %	Ablehnung in gültige %	M	SD
1	25	2	8,7	91,3	0,09	0,288
2	11	1	20,0	80,0	0,20	0,422
3	18	0	16,7	83,3	0,17	0,383
4	18	1	11,8	88,2	0,12	0,332
5	22	0	31,8	68,2	0,32	0,477
6	26	1	28,0	72,0	0,28	0,458

Tab. 46: Ankreuzverhalten der Klassen 1 – 6 hinsichtlich Aufgabe II.3, Betrachtung des Distraktors „Häufiges Kauen trainiert die Zähne“: Klassengröße N, Häufigkeit der Zustimmung und Ablehnung in Prozentangaben, Mittelwert und Standardabweichung.

Klasse	N	fehlend	Zustimmung in gültige %	Ablehnung in gültige %	M	SD
1	25	0	40,0	60,0	0,40	0,500
2	11	0	54,5	45,5	0,55	0,522
3	18	0	50,0	50,0	0,50	0,514
4	18	0	44,4	55,6	0,44	0,511
5	22	0	36,4	63,6	0,36	0,492
6	26	0	50,0	50,0	0,50	0,510

Tab. 47: Ankreuzverhalten der Klassen 1 – 6 hinsichtlich Aufgabe II.4, Betrachtung des Distraktors „Diese Studie beweist, dass Kaugummikauen keine Wirkung auf die Zähne hat“: Klassengröße N, Häufigkeit der Zustimmung und Ablehnung in Prozentangaben, Mittelwert und Standardabweichung.

Klasse	N	fehlend	Zustimmung in gültige %	Ablehnung in gültige %	M	SD
1	25	0	12,0	88,0	0,12	0,332
2	11	2	11,1	88,9	0,11	0,333
3	18	0	0,0	100,0	0,00	0,000
4	18	3	20,0	80,0	0,20	0,414
5	22	2	10,0	90,0	0,10	0,308
6	26	5	57,1	42,9	0,57	0,507

Tab. 48: Ankreuzverhalten der Klassen 1 – 6 hinsichtlich Aufgabe III.4, Betrachtung des Distraktors „Phenolphthalein“: Klassengröße N, Häufigkeit der Zustimmung und Ablehnung in Prozentangaben, Mittelwert und Standardabweichung.

E Schülerfragebogen der Hauptstudie

Liebe Schülerin, lieber Schüler!

In diesem Test findest du eine Reihe von Aufgaben und Fragen zu unterschiedlichen Themen. Einige Aufgaben sind kurz und andere sind etwas länger, ein paar Aufgaben werden dir schwerer fallen und andere werden leicht sein, einige haben viel, andere auf den ersten Blick weniger mit Chemie zu tun. Bitte versuche, alle Aufgaben so gut wie möglich zu lösen.

Dir werden verschiedene Typen von Aufgaben begegnen:

Bei einigen Aufgaben musst du deine Antwort zeichnen oder als kurzen Text oder Rechnung aufschreiben. Wenn der Platz auf dem Zettel nicht ausreicht, nimm bitte die Rückseite!

Bei anderen Aufgaben kreise bitte die richtigen Lösungen ein.

Fragen zur Person sind unverändert, vgl. Schülerfragebogen der Pilotstudie (Anhang A)

Sauer macht lustig?! Säuren machen Spaß?!

Nenne vier Stichworte, die dir aus deinem Chemieunterricht zum Thema „Säuren“ einfallen!

a) _____ b) _____

c) _____ d) _____

I. Was tun, wenn's im Magen brennt?

Dein Magen enthält eine saure Lösung, die die Verdauung unterstützt und Bakterien, die mit dem Essen in den Magen gelangen, abtötet. Wenn im Magen zu viel Säure produziert wird, dann „brennt“ es – z.B. in der Speiseröhre, wenn man Aufstoßen muss. Dies wird als „Sodbrennen“ bezeichnet.



I.1. Wenn du eine Tablette gegen die Übersäuerung des Magens herstellen solltest, **bei welchen Stoffen würdest du nach einer geeigneten Substanz suchen?**

- A Säuren
- B Laugen
- C sauer wirkende Salze
- D basisch wirkende Salze
- E Aktivkohle

I.2. Ganz unterschiedliche Aspekte müssen bei der Entwicklung eines neuen Medikaments berücksichtigt werden. **Nenne fünf Kriterien, die dir bei der Entwicklung eines Medikaments wichtig erscheinen!**

a) _____ b) _____ c) _____

d) _____ e) _____

I.3. Du möchtest dich über ein vorliegendes Medikament gegen Sodbrennen genauer informieren. Dazu hast du Gelegenheit, die folgenden Personen zu befragen.

Formuliere je eine Frage zu dem Medikament an

- deine Chemielehrerin / deinen Chemielehrer:
-

- deine Biologielehrerin / deinen Biologielehrer:
-

- eine Apothekerin / einen Apotheker:
-

Eine große Fülle von Informationen liefert auch das Internet.

Mit welchen drei Suchbegriffen startest du hier deine Recherche?

a) _____ b) _____ c) _____

I.4. Nach der Einnahme von Bullrich-Salz[®] muss man häufig aufstoßen. Grund ist die folgende chemische Reaktion im Magen: Der Wirkstoff Nat-



Bullrich-Salz[®] ist der Klassiker unter den „Magensäurestoppnern“!

Anwendung: Bei Sodbrennen, Magendruck, Völlegefühl.

Inhaltsstoffe: Natriumhydrogencarbonat, Talkum, Maisstärke

riumhydrogencarbonat bildet mit der im Magensaft enthaltenen Salzsäure Kohlenstoffdioxid.

Überlege dir einen Modellversuch, mit dem du verdeutlichen kannst, was mit einer Bullrich-Salz[®]-Tablette im Magen passiert und warum man nach ihrer Einnahme aufstoßen muss. Das entstehende Gas soll aufgefangen werden, um damit weitere Versuche durchführen zu können.

Erstelle eine Zeichnung der Versuchsanordnung und beschrifte sie ausführlich!

(Platz zum Zeichnen hast du auf der folgenden Seite!)

Versuchszeichnung:

- I.5. Der Packungsbeilage von Bullrich-Salz[®] ist zu entnehmen, dass in einer Tablette 840 mg Natriumhydrogencarbonat enthalten sind.

Formuliere die Reaktionsgleichung für die Neutralisationsreaktion von NaHCO_3 mit HCl !

Für die Wirksamkeit einer Tablette muss man wissen, wieviel Wirkstoff enthalten ist. **Gib an, wieviel Mol 840mg NaHCO_3 (enthalten in einer Tablette) entsprechen!**

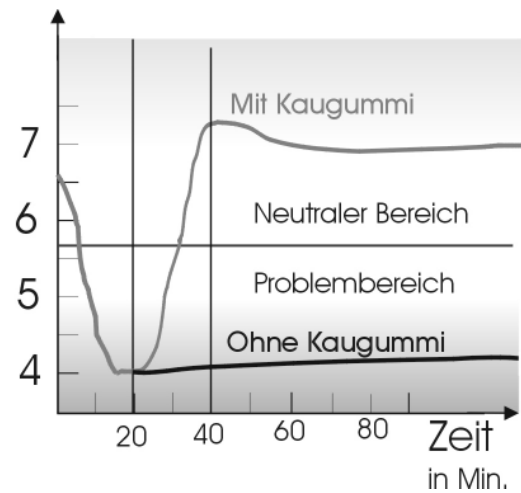
Hinweis: Bei der Berechnung der molaren Masse hilft dir das abgebildete Periodensystem!

1 H 1,0079								2 He 4,0026
3 Li 6,94	4 Be 9,0122							
11 Na 22,990	12 Mg 24,305	5 B 10,811	6 C 12,011	7 N 14,007	8 O 15,999	9 F 18,998	10 Ne 20,180	
19 K 39,098	20 Ca 40,078	13 Al 26,982	14 Si 28,086	15 P 30,974	16 S 32,066	17 Cl 35,453	18 Ar 39,948	
37 Rb	38 Sr	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr	

Was musst du in Erfahrung bringen, um zu entscheiden, wie viele Tabletten ein Patient mit Sodbrennen einnehmen sollte.

II. Kaugummi gegen Karies

Nach der Nahrungsaufnahme entstehen durch Bakterien im Mund Säuren. Diese Säuren schädigen den Zahnschmelz, die Zähne werden anfällig für Karies. Eine Kaugummi-Werbung verspricht, durch ein speziell entwickeltes Zahnpflege-Kaugummi das Problem der Kariesanfälligkeit der Zähne nach dem Essen zu lösen. Mit der nebenstehenden Grafik wirbt das Unternehmen für seine Kaugummis. Der Liste der Inhaltsstoffe auf der Kaugummi-Verpackung ist zu entnehmen, dass in diesen Kaugummis kein basisch reagierender Stoff enthalten ist.



II.1. Sieh dir die abgebildete Grafik an! **Beschreibe die Veränderungen des pH-Werts und gehe dabei insbesondere auf den Verlauf der Kurve nach 20, 40 und 60 Minuten ein!**

Was passiert nach Aussage des Kaugummi-Herstellers, wenn du nach dem Essen eines seiner Kaugummis kaust?

II.2. Sieh dir die Abbildung noch einmal an: Solche Bilder sollen vor allem werbewirksam sein. Manchmal lohnt es sich aber, ein bisschen genauer hinzuschauen.

Worüber stolpert dein „chemischer Sachverstand“ in dieser Grafik?

II.3. Geht deinen Eltern „Kaugummikauen“ auch auf die Nerven?

Mit welchen der folgenden Argumente kannst du erklären, dass du deinen Zähnen mit dem „Kaugummikauen“ etwas Gutes tust? Kennzeichne „ja“ oder „nein“!

Ist das ein Argument zu Gunsten des „Kaugummikauens“?	Ja oder nein?
Zucker im Kaugummi konserviert die Zähne.	ja / nein
Durch Kauen angeregter Speichelfluss reinigt die Zähne.	ja / nein
Das Kauen eines Kaugummis reinigt die Zähne.	ja / nein
Häufiges Kauen trainiert die Zähne.	ja / nein

II.4. In einer Studie wurden 225 Personen gebeten, drei Monate lang täglich drei Mal ein Zahnpflege-Kaugummi zu kauen. Die Testpersonen wurden vor und nach dieser Zeit zahnärztlich untersucht. Bei 175 Personen war der Zustand unverändert, bei 40 hatte er sich leicht verbessert, und bei den restlichen 10 geringfügig verschlechtert.

Welche der folgenden Aussagen ist korrekt?

- A Diese Studie beweist, dass Kaugummikauen gut für die Zähne ist.
- B Diese Studie beweist, dass Kaugummikauen keine Wirkung auf die Zähne hat.
- C Diese Studie beweist, dass Kaugummikauen schlecht für die Zähne ist.
- D Diese Studie beweist gar nichts.



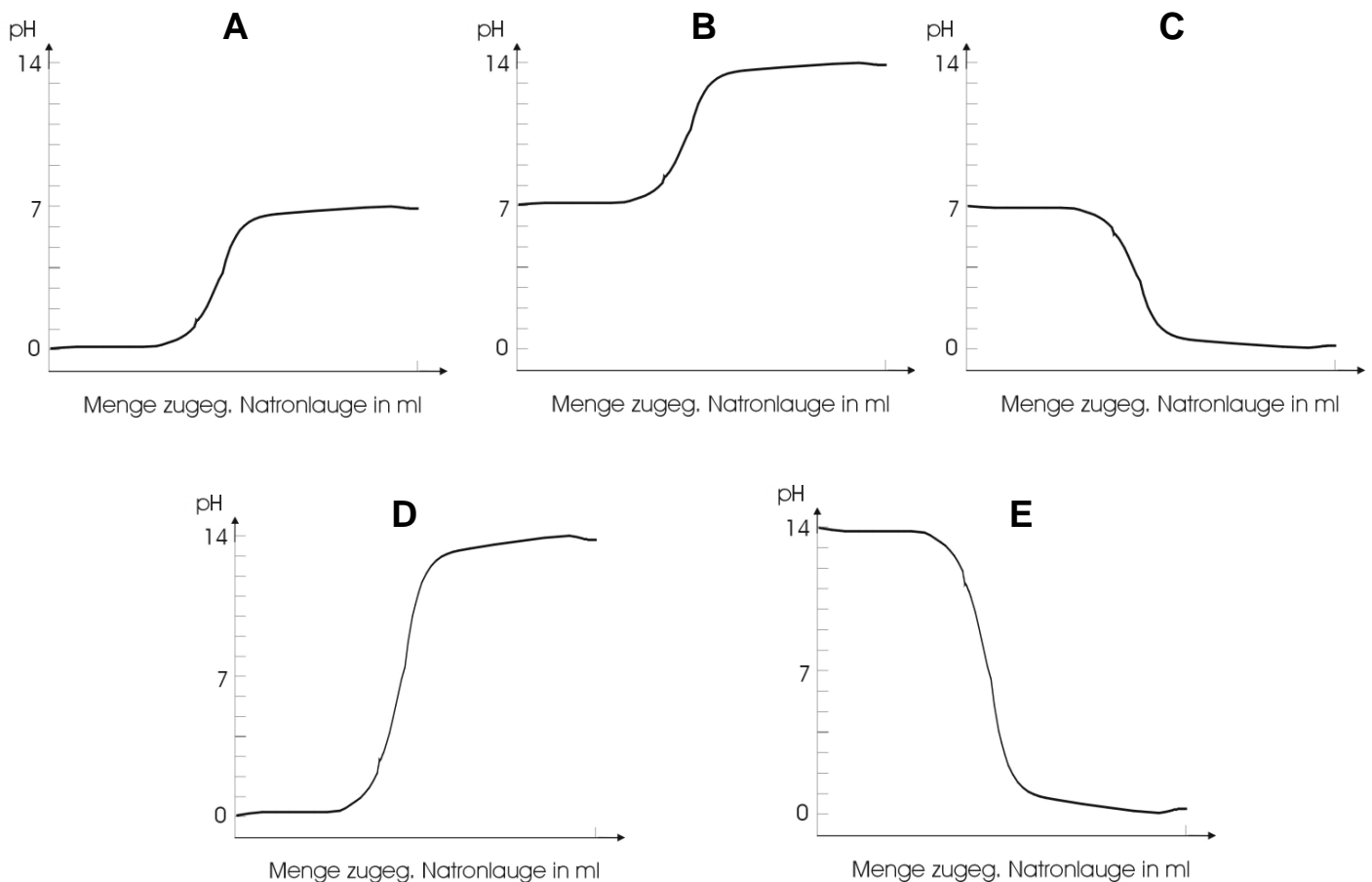
**BITTE UMBLÄTTERN –
JETZT KOMMT DER
LETZTE TEIL DES
TESTS!!**

III. Im Chemielabor

III.1. Der Chemikalienschrank ist gut sortiert. In welcher Aufzählung werden ausschließlich Stoffe genannt, die in Wasser gelöst basisch (alkalisch) reagieren?

- A NaOH, CH₃COOH, H₂SO₄
- B NaOH, CaO, NaHCO₃
- C NaCl, NH₃, KOH
- D CH₃CH₂OH, CaO, HCl

III.2. In einem Gefäß befindet sich Salzsäure. Zu der Salzsäure wird tropfenweise Natronlauge zugegeben. Diese Natronlauge besitzt die gleiche Konzentration wie die Salzsäure, ist aber in der doppelten Menge vorhanden. Nach und nach wird die gesamte Natronlauge zugetropft: **Kennzeichne, welche der folgenden Kurven die pH-Wert-Änderung richtig wiedergibt!**



- III.3.** Beim Experimentieren wurde im Abzug Säure verschüttet. Eine Mitschülerin schlägt vor, Wasser hinzuzugeben, ein anderer Mitschüler möchte Natronlauge nehmen.
Welchem Ratschlag schließt du dich an? Begründe deine Entscheidung!

- III.4.** Indikatoren verändern je nach pH-Wert einer Lösung ihre Farbe. **Kennzeichne, mit welchem der drei Indikatoren du prüfen kannst, ob Regenwasser pH-neutral ist!**

A	Phenolphthalein	Umschlagbereich bei pH 8,0 (farblos) – 9,8 (violett)
B	Methylorange	Umschlagbereich bei pH 3,1 (orange) – 4,4 (gelb)
C	Bromthymolblau	Umschlagbereich bei pH 6,2 (gelb) – 7,6 (blau)

- III.5.** In der linken Tabellenspalte sollen Eigenschaften eingetragen werden, die allen sauren Lösungen gemeinsam sind. In der rechten Tabellenspalte soll dazu je ein Versuch eingetragen werden, mit dem die links genannte Eigenschaft demonstriert werden kann.

Ergänze alle Felder der Tabelle!

Eigenschaft / Merkmal einer sauren Lösung	geeigneter Versuch zur Demonstration

- III.6.** Destilliertes Wasser leitet den elektrischen Strom praktisch nicht.
Eine saure Lösung hingegen leitet den elektrischen Strom gut.

**Erläutere bitte das Phänomen mit Hilfe der Teilchen des destillierten Wassers
und der sauren Lösung!**

**VIELEN DANK FÜR DEINE
MITARBEIT UND DEIN
DURCHHALTEVERMÖGEN!!**



F Erwartungshorizont zur Bewertung der Aufgaben mit offenem Antwortformat (Hauptstudie)

Aufgabe	Erwartungshorizont (für mehrstufiger Auswertung wird eine Stufung Antworten angegeben)
I.2	Für jedes der fünf angegebenen Kriterien wird entscheiden, ob es in einem sinnvollen Zusammenhang zu einem Medikament steht, als „sinnvoll“ wird angesehen: Preis, Wirtschaftlichkeit, einfache, kostengünstige Herstellung, keine Giftwirkung, Nebenwirkungen, Verträglichkeit, Wirkung, Geschmack, einfache Aufnahme/Einnahme, dabei werden „Verträglichkeit“, „schädliche Wirkung auf den Körper“ und „Nebenwirkungen“ als drei richtige Kriterien gewertet.
I.3	Folgende Fragen werden als richtig bewertet: <i>Chemielehrerin/Chemielehrer:</i> z.B. Welche Stoffe sind in dem Medikament enthalten? Wie wirkt das Medikament? Welche chemischen Eigenschaften hat der Wirkstoff? Welche chemische Struktur hat der Wirkstoff/das Medikament? Welche Summenformel hat der Wirkstoff? Welche chemischen Reaktionen laufen im Körper ab? Welcher pH-Wert liegt im Magen vor? Wie kann man die Wirkungsweise chemisch erklären? <i>Biologielehrerin/Biologielehrer:</i> z.B. Warum ist der pH-Wert im Magen so niedrig? Wozu brauchen wir die Säure im Magen? Was passiert mit dem Medikament im Magen/Körper? Wie wird das Medikament abgebaut? Wie wirkt das Medikament im Körper? (Fragen zum Abbau im Körper, Weg im Körper, biologischen Ursprung des Medikaments) <i>Apothekerin/Apotheker:</i> z.B. Wie wirkt das Medikament? Welche Nebenwirkungen hat das Medikament? Würden Sie das Medikament empfehlen? Welche Medikamente gegen Sodbrennen gibt es? Was kosten die verschiedenen Medikamente? Was empfehlen Sie mir? Welche Wechselwirkungen können auftreten? Wann darf man das Medikament nicht einnehmen?
I.3b	Es werden alle Suchbegriffe als richtig gewertet, wenn sie einen Bezug zur Krankheit oder zum Medikament haben, z.B. Sodbrennen, Magensäure, Medikamente, Bullrich Salz, Neutralisation, Aufstoßen, Nebenwirkungen, Hersteller, Name des Medikaments, ...; die Krankheit muss in irgendeiner Weise als Suchbegriff auftauchen, einzelne Kombinationen sind nicht sinnvoll, wenn die Begriffe zwar richtig sind, aber keinen Bezug zum Medikament oder zur Krankheit herstellen
I.4	Bewertung der Versuchszeichnung: Sechs Zeichnungselemente werden nach ihrem Vorhandensein bewertet, die Kombination der sechs Ziffern liefert dann eine Gesamtbewertung. Dabei gilt übergeordnet: eine Versuchsapparatur gilt als korrekt, wenn sie nur deshalb nicht funktioniert, weil unberücksichtigt blieb, dass Kohlenstoffdioxid schwerer als Luft ist, wenn der Mensch in der Zeichnung das Reaktionsgefäß darstellt, wenn ein Luftballon oder eine Tüte als Auffangbehältnis des Gases fungieren. 3 Punkte: Die Zeichnung gibt einen sinnvollen Versuchsaufbau wieder. Aus der Zeichnung wird deutlich, wie das Gas aufgefangen (z.B. Luftballon) oder abgeleitet wird und worin sich das Reaktionsgemisch befindet. Das Kohlenstoffdioxid muss nicht rein aufgefangen werden, für Nachweise reicht ein Gemisch mit Luft. Darüber hinaus wird unterstellt, dass die Lernenden wissen, dass das aufgefangene Gas schwerer als Luft ist: Ein nach oben offenes Gefäß funktioniert als Auffanggefäß für das entstehende Gas. Das Gas wird so aufgefangen, dass damit weitere Versuche durchgeführt werden können (Glimmspan, Kerze, Barytwasser, es geht nicht um eine quantitative Bestimmung). Die eingesetzten Gefäße/Geräte und die Chemikalien sind beschriftet. 2 Punkt: insgesamt eine wie oben beschriebene Darstellung einer offenen Versuchsapparatur mit einer Schwäche, z.B. unvollständige/fehlende Beschriftung, Einleitung des Gases in eine Nachweislösung oder unpräzise Zeichnung. 1 Punkt: Die Zeichnung gibt insgesamt den Versuchsaufbau richtig, aber als ein geschlossenes System wieder. 0 Punkte: Die Zeichnung ist grundsätzlich fehlerhaft.
I.5.a	Die Reaktionsgleichung muss vollständig und stöchiometrisch korrekt sein. Eine Darstellung des Reaktionspfeils als Gleichheitszeichen wird nicht als Fehler gewertet, nicht zerfallene Kohlensäure wird ebenso wie undissoziiertes Natriumchlorid nicht als Fehler gewertet.
I.5.b	Für die Berechnung wird festgelegt: 2 Punkte: Die Berechnung wird vollständig und korrekt durchgeführt (Ergebnis: 0,01 mol). 1 Punkt: Die Molmasse von NaHCO_3 wird korrekt bestimmt und mit richtigen Einheiten angegeben (84g/mol). 0 Punkte: falsche Ergebnisse, fehlende/falsche Einheiten bei richtigem Zahlenwert
I.5.c	2 Punkte: Es werden eine oder mehrere sinnvolle Informationen erfragt, die einen Bezug zur Menge des Überschusses der sauren Lösung im Magen haben, dabei werden Formulierungen wie z.B. „Höhe des Übersäuerung“, „wie viel Säure neutralisiert werden muss“ oder „wie stark die Übersäuerung ist“ gewertet. 1 Punkt: Es werden eine oder mehrere sinnvolle Informationen erfragt, dabei aber kein Bezug zur Menge des Überschusses genommen. 0 Punkte: Es werden Informationen erfragt, die nicht in einem sinnvollen Zusammenhang mit der Krankheit stehen. (Als sinnvolle Informationen gelten z.B. „Konzentration der Magensäure, Alter, Gewicht, Essgewohnheiten, Stärke des Sodbrennens, Dosierungsanleitung, ...)
II.1.a	Als richtig werden korrekte Beschreibungen des Kurvenverlaufs bewertet, die einen Bezug zur Grafik herstellen, z.B. durch Nennung von Zeitangaben, Angabe von pH-Werten. Falsche Markierungen in der Kurve werden nicht bewertet, können aber Aufschluss über das Verständnis der Kurve geben.
II.1.b	2 Punkte: Für eine richtige Interpretation werden Informationen aus der Grafik entnommen und korrekt gedeutet. 1 Punkt: Die Informationen für die Interpretation werden in erster Linie aus dem Aufgabentext entnommen, die Interpretation ist korrekt, bezieht sich aber nicht auf die Grafik. 0 Punkte: aus der richtigen Interpretation werden falsche Schlussfolgerungen gezogen.
II.2.	Für eine positive Bewertung wird die Nennung eines Aspektes, z.B. unterschiedliche Begriffsdefinition von pH-neutral in Chemie und Werbung; Zweifel an der dargestellten Information, dass der pH-Wert im Mund über einen recht langen Zeitraum bei ca. 4 liegt; Lage des pH-Werts vor dem Essen bei 6,5, nach dem Kauen des Kaugummis bei ca. 7, fehlende Zeitangabe.
III.3	2 Punkte: Entscheidet sich nicht für Natronlauge, da man zuvor genau die Menge Natronlauge berechnen müsste, die man benötigt, um die Säure zu neutralisieren. Entscheidet sich für (viel) Wasser, um die Säure (stark) zu verdünnen (und damit pH-Wert zu erhöhen). Entscheidet sich für Natronlauge, verweist aber darauf, dass der pH-Wert kontrolliert werden muss, damit man nicht zuviel Natronlauge zugibt. [Ein Verweis auf die freiwerdende Neutralisationswärme als zusätzliche Gefahrenquelle ist nicht notwendig zur vollständigen und korrekten Lösung dieser Aufgabe.] 1 Punkt: Entscheidet sich für Natronlauge, weil es die Säure neutralisiert. 0 Punkte: Schließt sich keinem Ratschlag an oder begründet Entscheidung nicht oder falsch.

Anhang

Aufgabe	Erwartungshorizont (für mehrstufiger Auswertung wird eine Stufung Antworten angegeben)								
III.5.	<table border="1"> <thead> <tr> <th>Eigenschaft/Merkmal</th> <th>geeigneter Versuch zur Demonstration</th> </tr> </thead> <tbody> <tr> <td>saure Lösungen leiten den elektrischen Strom</td> <td>Durchführung einer Leitfähigkeitsmessung</td> </tr> <tr> <td>ätzende Wirkung auf (unedle) Metalle, Freisetzung von Wasserstoff bei der Reaktion mit (unedlen) Metallen</td> <td>Überprüfung der Reaktion mit (unedlen) Metallen, Untersuchung der Gasbildung, Knallgasprobe mit freigesetztem Gas</td> </tr> <tr> <td>pH-Wert < 7; färbt Indikatoren charakteristisch; neutralisieren basische Lösungen, greifen Kalk an</td> <td>Nachweis mit pH-Papier/ Indikatorlösungen/ pH-Wert-Messung; Gasbildung mit Kalk</td> </tr> </tbody> </table> <p>3Punkte: Füllt alle Felder der Tabelle korrekt aus. 2 Punkte: Nennt zwei richtige Eigenschaften, davon eine mit geeignetem Versuch. 1 Punkt: Nennt eine richtige Eigenschaft mit einem geeigneten Versuch oder zwei richtige Eigenschaften.</p>	Eigenschaft/Merkmal	geeigneter Versuch zur Demonstration	saure Lösungen leiten den elektrischen Strom	Durchführung einer Leitfähigkeitsmessung	ätzende Wirkung auf (unedle) Metalle, Freisetzung von Wasserstoff bei der Reaktion mit (unedlen) Metallen	Überprüfung der Reaktion mit (unedlen) Metallen, Untersuchung der Gasbildung, Knallgasprobe mit freigesetztem Gas	pH-Wert < 7; färbt Indikatoren charakteristisch; neutralisieren basische Lösungen, greifen Kalk an	Nachweis mit pH-Papier/ Indikatorlösungen/ pH-Wert-Messung; Gasbildung mit Kalk
Eigenschaft/Merkmal	geeigneter Versuch zur Demonstration								
saure Lösungen leiten den elektrischen Strom	Durchführung einer Leitfähigkeitsmessung								
ätzende Wirkung auf (unedle) Metalle, Freisetzung von Wasserstoff bei der Reaktion mit (unedlen) Metallen	Überprüfung der Reaktion mit (unedlen) Metallen, Untersuchung der Gasbildung, Knallgasprobe mit freigesetztem Gas								
pH-Wert < 7; färbt Indikatoren charakteristisch; neutralisieren basische Lösungen, greifen Kalk an	Nachweis mit pH-Papier/ Indikatorlösungen/ pH-Wert-Messung; Gasbildung mit Kalk								
III.6	<p>2 Punkte: Nennt die Existenz von Ionen in der Lösung als Voraussetzung für elektrische Leitfähigkeit und erklärt die Leitfähigkeit von sauren Lösungen über Bildung von H_3O^+-Ionen/über die Bildung von Protonen und Säure-Anionen (Säuren sind Protonendonatoren)/über Protonen/Ladungsträger/positiv oder negativ geladene Teilchen. In der Erläuterung wird sowohl auf das destillierte Wasser als auch auf die saure Lösung eingegangen.</p> <p>1 Punkt: Die Begründung ist nicht vollständig, aber korrekt: z.B. „durch vorliegende Ionen kann Strom geleitet werden“ oder „Säure bildet in Wasser Ionen“.</p>								

Tab. 49: Erwartungshorizont zur Bewertung der Aufgaben mit offenem Antwortformat der Hauptstudie, bei mehrstufigen Items ist für jede Stufung die erforderliche Bearbeitung beschrieben.

G Kurzbeschreibung der Testaufgaben (Hauptstudie)

Aufgabe	Kurzbezeichnung	Kurzbeschreibung
I.1	„Gegenmittel Sodbrennen“	Aus einer Liste von fünf Substanzen müssen die Probanden die Substanz auswählen, die ihnen zur Bekämpfung von Magenübersäuerung geeignet erscheint.
I.2	„Aspekte der Medikamententwicklung“	Die Probanden sind aufgefordert, fünf Aspekte zu nennen, die ihnen bei der Entwicklung eines neuen Medikamentes wichtig erscheinen.
I.3	„Fragen formulieren“	Die Probanden sind aufgefordert, je eine Frage zu formulieren, die sie ihrer Chemie- und Biologielehrkraft und einem Apotheker/einer Apothekerin stellen würden, um sich genauer über ein Medikament gegen Sodbrennen zu informieren.
	„Internet-Suchbegriffe“	Die Probanden sind aufgefordert, drei Suchbegriffe zu formulieren, mit denen sie eine Internetrecherche zu einem Medikament gegen Sodbrennen beginnen würden.
I.4	„Versuch zum Gas auffangen“	Die Probanden sind aufgefordert, eine Versuchszeichnung zu einem Demonstrationsversuch zu erstellen, mit dem sie das entstehende Gas bei der Reaktion von NaHCO_3 mit saurer Lösung auffangen könnten.
I.5	„Reaktionsgleichung“	Die Probanden sind aufgefordert, die Reaktionsgleichung für die Neutralisationsreaktion von NaHCO_3 mit HCl zu formulieren.
	„Berechnung“	Die Probanden sind aufgefordert zu berechnen, wie viel Mol 840 mg NaHCO_3 entsprechen.
	„Informationen einholen“	Die Probanden sind aufgefordert, sich zu überlegen, welche Information(en) sie benötigen, um zu entscheiden, wie hoch die Einnahme eines Medikamentes gegen Magenübersäuerung dosiert werden muss.
II.1	„Grafik beschreiben“	Die Probanden sind aufgefordert, eine graphische Darstellung zur pH-Wert-Veränderung im Mund mit und ohne Kauen eines Zahnpflekekaugummi mit eigenen Worten zu beschreiben.
	„Grafik interpretieren“	Die Probanden sind aufgefordert, die graphische Darstellung in Hinblick auf die Wirkungsweise eines Zahnpflekekaugummi zu interpretieren.
II.2	„Fehlersuche“	Die Probanden sind aufgefordert, die graphische Darstellung auf ihre Richtigkeit hin zu überprüfen und Fehler in der Darstellung zu identifizieren.
II.3	„Kaugummiwirkung“	Die Probanden sind aufgefordert, für vier Aussagen zur Wirkungsweise von Zahnpflekekaugummi jeweils zu entscheiden, ob die Aussage ein Argument für das Kaugummikauen darstellt.
II.4	„Studie bewerten“	Die Probanden sind aufgefordert, die Aussage einer beschriebenen Studie zur Wirksamkeit von Zahnpflekekaugummi zu bewerten, indem sie sich für eine von vier dargestellten möglichen Aussagen der Studie entscheiden.
III.1	„basische Stoffe“	Aus vier verschiedenen Aufzählungen von je drei Stoffen müssen die Probanden die Aufzählung auswählen, in der ausschließlich Stoffe genannt werden, die in Wasser gelöst basisch reagieren.
III.2	„pH-Kurven“	Aus fünf verschiedenen pH-Wert-Auftragungen müssen die Probanden die Darstellung auswählen, die die pH-Wert-Änderung bei einer Titration einer Salzsäurelösung mit einer gleich konzentrierten Natronlauge korrekt wiedergibt.
III.3	„Ratschlag“	Die Probanden sind aufgefordert, sich einem der beiden Ratschläge anzuschließen und diese Entscheidung zu begründen, wie mit verschütteter Säurelösung umgegangen werden sollte.
III.4	„Indikatoren“	Aus einer Liste von drei Indikatoren und deren pH-Umschlagsbereichen müssen die Probanden den Indikator auswählen, mit dem die pH-Neutralität von Regenwasser überprüft werden kann.
III.5	„Demonstrationsversuche“	Die Probanden sind aufgefordert, in tabellarischer Form zwei Eigenschaften von sauren Lösungen zu nennen und für jede der beiden Eigenschaften einen Versuch anzugeben, mit dem die genannte Eigenschaft demonstriert werden kann.
III.6	„Leitfähigkeit“	Die Probanden sind aufgefordert, mit Hilfe der Teilchenebene zu erklären, warum saure Lösungen den elektrischen Strom gut leiten, destilliertes Wasser hingegen den elektrischen Strom praktisch nicht leitet.

Tab. 50: Kurzbeschreibung und Kurzbezeichnung der eingesetzten Testaufgaben aus der Hauptstudie.

H Lehrerfragebogen der Hauptstudie

Liebe Kollegin, lieber Kollege!

Sie haben sich bereit erklärt, in Ihrer Chemiekunde die von uns entwickelten Aufgaben zum Thema „Sodbrennen, Kaugummi und Chemie“ einzusetzen. Dafür bedanken wir uns sehr herzlich!

Vermerken Sie bitte zunächst Ihren **Projekt-Code**

3. Buchstabe des Vornamens des Vaters	2. Buchstabe des Vornamens der Mutter	1. Buchstabe d. Mädchennamens der Mutter	3. Buchstabe des eigenen Geburtsortes	Ihr Geburtsmonat zweistellig (z.B. April = 04)

Um die Ergebnisse der einzelnen Schüler/-innen bzw. Klassen vergleichen zu können, benötigen wir Informationen über den von Ihnen in der befragten Klasse durchgeführten Unterricht zum Thema „Säuren und Basen“ und zu den Erfahrungen mit ChiK-Unterricht.

Haben Sie das „Säure-Base-Konzept“ in der befragten Klasse im Rahmen der Unterrichtskonzeption „Chemie im Kontext“ behandelt? ja nein

Wenn **ja**, in welchem **Kontext** haben Sie das Thema „Säuren und Basen“ unterrichtet?

Wenn **nein**: Bei dem Thema kann es sich anbieten, Bezüge zum Alltag herzustellen, auf welche „**Alltagsthemen**“ haben Sie ggf. Bezug genommen?

Wie viele **Unterrichtsstunden** umfasste die Einheit zum Thema „Säuren und Basen“?

Ca. _____ Stunden

Wann haben Sie die Einheit unterrichtet (Monat / Jahr)? _____

Wie viele **ChiK-Einheiten** haben Sie bereits insgesamt unterrichtet?

_____ ChiK-Einheiten

Welche **Kontexte**?

Haben die befragten Schüler/innen nach Ihrem Kenntnisstand Erfahrungen mit ChiK-Unterricht? ja nein

Falls ja, wie viele ChiK-Einheiten sind in dieser Klasse unterrichtet worden?

_____ ChiK-Einheiten.

Welche Themen (Kontexte)?

Auf den folgenden Seiten finden Sie den von uns formulierten **Erwartungshorizont** zu den einzelnen Aufgaben und eine kurze Darstellung, welche **Anforderung** wir mit der jeweiligen Aufgabe insbesondere erfassen wollen. Auf diese Weise möchten wir u.a. genauer erfahren, ob die einzelnen Aufgabeninhalte Thema in dem Unterricht der befragten Klasse waren!

Wir möchten Sie bitten, uns zu jeder Aufgabe durch ein Kreuz **Ihre Einschätzung** mitzuteilen, wieviel Prozent der Schüler/innen der befragten Klasse die Aufgabe lösen können, wie interessant Ihre Schüler die Art der Aufgabe finden, ob die Anforderung einen Inhalt Ihres Unterrichts in der befragten Klasse darstellte und ob Sie Aufgaben dieser Art in Ihrem Unterricht einsetzen.

Auf den folgenden Seiten sind alle Teilaufgaben wie im Beispiel gezeigt dargestellt:

1. Aufgabentext (ggf. gekürzt)
2. die erwartete korrekte Antwort (kursiver bzw. eingekreist)
3. Anforderung der Aufgabe¹ und
4. Tabelle für Ihre Einschätzungen

Beispiel:

(1. Aufgabentext): Welche Farbe hat das Trikot des neuen Fußball-Europameisters 2004?

- A rot-grün
- B orange-schwarz
- C blau-weiß
- D gelb-grün

(2. korrekte Antwort in kursiver Schrift bzw. eingekreist)

(2. Anforderung der Aufgabe):

Wissen: Sieger der EM 2004, Trikotfarbe der griechischen Fußball-Nationalmannschaft

(3. Tabelle für Ihre Einschätzung):

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
o eher ja <input checked="" type="radio"/> eher nein	0.....20.....40...50...60.....80.....100 <input checked="" type="radio"/>	0.....20.....40...50...60.....80.....100 <input checked="" type="radio"/>	o eher ja <input checked="" type="radio"/> eher nein

¹ Auf der letzten Seite dieses Fragebogens finden Sie eine Zusammenstellung einiger Anforderungskategorien mit kurzen Erläuterungen, um möglichen Begriffsunstimmigkeiten vorzubeugen!

ANTACIDA - WAS TUN, WENN'S IM MAGEN BRENNT?

I.1. Wenn du eine Tablette gegen die Übersäuerung des Magens herstellen solltest, bei welchen Stoffen würdest du nach einer geeigneten Substanz suchen?

- A Säuren
- B Laugen**
- C sauer wirkende Salze
- D basisch wirkende Salze**
- E Aktivkohle

Anforderung der Aufgabe

Wissen: Neutralisationsreaktion;

konvergentes Denken: Problem und Neutralisationsreaktion als Problemlösung

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
o eher ja o eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	o eher ja o eher nein

I.2. Ganz unterschiedliche Aspekte müssen bei der Entwicklung eines neuen Medikaments berücksichtigt werden. Nenne fünf Kriterien, die dir bei der Entwicklung eines Medikaments wichtig erscheinen!

alle fünf genannten Aspekte sollten in einem sinnvollen Zusammenhang zu Medikamenten stehen (z.B. Preis, Nebenwirkungen des Medikaments, Haltbarkeit, schnelle Wirkung, Geschmack, Darreichungsform, Verfügbarkeit, Werbung, ...)

Anforderung der Aufgabe

divergentes Denken: Faktoren zur Medikamentenentwicklung

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
o eher ja o eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	o eher ja o eher nein

I.3. Du möchtest dich über ein vorliegendes Medikament gegen Sodbrennen genauer informieren. Dazu hast du Gelegenheit, die folgenden Personen zu befragen. Formuliere je eine Frage zu dem Medikament an

- deine Chemielehrerin / deinen Chemielehrer:
z.B. Welche Stoffe sind in dem Medikament enthalten? Was passiert mit der Säure im Magen? Welche chemischen Eigenschaften hat der Wirkstoff? Welche chemische Struktur hat der Wirkstoff/das Medikament? Welche Summenformel hat der Wirkstoff? Welche chemischen Reaktionen laufen im Körper ab? ...
- deine Biologielehrerin / deinen Biologielehrer:
Warum ist der pH-Wert im Magen so niedrig? Wozu brauchen wir die Säure im Magen? Was passiert mit dem Medikament im Magen/Körper? Wie wird das Medikament im Körper abgebaut? Was macht die Säure im Magen?...
- eine Apothekerin / einen Apotheker:
Welche Nebenwirkungen hat das Medikament? Würden Sie das Medikament empfehlen? Welche Medikamente gegen Sodbrennen gibt es? Was kosten die verschiedenen Medikamente? ...

Eine große Fülle von Informationen liefert auch das Internet: Mit welchen drei Suchbegriffen startest du hier deine Recherche?

Sodbrennen, Magensäure, Medikamente, „Name des Medikaments“, Neutralisation, „Name des Herstellers“, Nebenwirkungen, ...

Anforderung der Aufgabe

divergentes Denken: Formulierung von Fragen (adressaten- / situationsgerecht);

konvergentes Denken: Planung einer (Internet-) Recherche

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

I.4. Überlege dir einen Versuch, mit dem du verdeutlichen kannst, was mit einer Bullrich-Salz[®]-Tablette im Magen passiert und warum man nach ihrer Einnahme aufstoßen muss. Das entstehende Gas soll aufgefangen werden, um damit weitere Versuche durchführen zu können. Erstelle eine Zeichnung der Versuchsanordnung und beschrifte sie ausführlich!

Die Zeichnung gibt einen sinnvollen, in der Praxis funktionierenden Versuchsaufbau wieder. Aus der Zeichnung wird deutlich, wie das Gas aufgefangen wird (z.B. Luftballon, Kolbenprober), wie das Gas abgeleitet wird, worin sich das Reaktionsgemisch befindet. Das Gas wird so aufgefangen, dass damit weitere Versuche durchgeführt werden können (es geht nicht um eine quantitative Bestimmung, das aufgefangene Gas muss nicht rein sein).

Die eingesetzten Gefäße/Geräte und die Chemikalien sind ausreichend beschriftet.

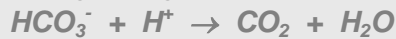
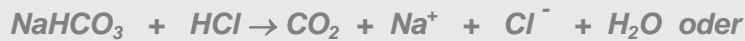
Anforderung der Aufgabe

konvergentes Denken: Planung eines Experiments auf der Basis der Fachinformationen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

I.5. Der Packungsbeilage von Bullrich-Salz[®] ist zu entnehmen, dass in einer Tablette 840 mg Natriumhydrogencarbonat enthalten sind.

Formuliere die Reaktionsgleichung für die Neutralisationsreaktion von NaHCO₃ mit HCl!



(ggf. jeweils Zwischenschritt über zerfallende „Kohlensäure“)

Anforderung der Aufgabe

Wissen: Neutralisation, Aufstellen einer Reaktionsgleichung

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

Gib an, wieviel Mol 840mg NaHCO₃ (enthalten in einer Tablette) entsprechen!

Bestimmung der molaren Masse von NaHCO₃: $M(\text{NaHCO}_3) = 84\text{g/mol}$;

840 mg NaHCO₃ entsprechen 0,01 Mol NaHCO₃

Anforderung der Aufgabe

Wissen: Begriff der „molaren Masse“;

Umgang mit Zahlen: stöchiometrische Berechnungen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

Was musst du in Erfahrungen bringen, um zu entscheiden, wie viele Tabletten ein Patient mit Sodbrennen einnehmen sollte.

z.B. Die überschüssige Säuremenge im Magen. Menge und Konzentration der sauren Lösung im Magen. ...

auch: Die Dosierungsanleitung aus der Packungsbeilage. ...

Anforderung der Aufgabe

Wissen: Neutralisation;

konvergentes Denken: Informationen zusammenbringen, um eine Entscheidung treffen zu können

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

KAUGUMMI GEGEN KARIES

- II.1.** Sieh dir die abgebildete Grafik an! Beschreibe die Veränderungen des pH-Werts und gehe dabei insbesondere auf den Verlauf der Kurve nach 20, 40 und 50 Minuten!

Der Text sollte eine Beschreibung der Kurve wiedergeben, dabei soll auf die genannten Zeitpunkte eingegangen werden. (z.B. Nach dem Essen liegt der pH-Wert im Mund bei 4 /im sauren Bereich, nach 20 Minuten steigt der pH-Wert/die Kurve auf über 7, bei 40 Minuten ist sie am höchsten (bei ca. 7,3), dann fällt die Kurve wieder etwas und bleibt dann nach 50 Minuten etwa konstant bei pH=7 / im pH-neutralen Bereich.)

Was passiert nach Aussage des Kaugummi-Herstellers, wenn du nach dem Essen eines seiner Kaugummis kaust!

z.B. Durch das Kauen des Kaugummis werden die Säuren im Mund neutralisiert. Das Kaugummi neutralisiert den Speichel. Die Säuren im Mund werden durch das Kauen des Kaugummis unschädlich gemacht. ...

Anforderung der Aufgabe

Wissen: pH-Wert;

Umgang mit grafischen Darstellungen: Verbalisierung grafisch dargestellter Informationen, Interpretation grafisch dargestellter Informationen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

- II.2.** Sieh dir die Abbildung noch einmal an: Solche Bilder sollen vor allem werbewirksam sein. Manchmal lohnt es sich aber, ein bißchen genauer hinzuschauen. Worüber stolpert dein „chemischer Sachverstand“ in dieser Grafik?

Unterschiedliche Begriffsdefinition: In der Chemie versteht man unter „neutral“ eigentlich „pH-neutral“, und damit pH 7, in dieser Abbildung ist der Begriff „neutral“ um pH 6 vermerkt und bezieht sich nicht auf die Definition aus der Chemie. Ebenfalls richtig: Zweifel an der dargestellten Information, dass der pH-Wert im Mund über einen recht langen Zeitraum unverändert bei ca. 4 liegt, wenn man kein Kaugummi kaut; pH-Wert liegt vorher bei 6,5, nach dem Kauen des Kaugummis darüber bei ca. 7, Zweifel daran, dass der pH-Wert im Mund zeitweise sogar über 7 liegen soll

Anforderung der Aufgabe

Wissen: pH-Wert;

Umgang mit grafischen Darstellungen; Umgang mit Fachsprache; Bewertung von Informationen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

II.3. Geht deinen Eltern „Kaugummikauen“ auch auf die Nerven?
Mit welchen der folgenden Argumente kannst du erklären, dass du deinen Zähnen mit dem „Kaugummikauen“ etwas Gutes tust? Kennzeichne „ja“ oder „nein“!

Ist das ein Argument zu Gunsten des „Kaugummikauens“?	Ja oder nein?
Zucker im Kaugummi konserviert die Zähne.	ja / <input checked="" type="radio"/> nein
Durch Kauen angeregter Speichelfluss reinigt die Zähne.	<input checked="" type="radio"/> ja / nein
Das Kauen eines Kaugummis reinigt die Zähne.	<input checked="" type="radio"/> ja / nein
Häufiges Kauen trainiert die Zähne.	ja / <input checked="" type="radio"/> nein

Anforderung der Aufgabe

konvergentes Denken: Problem der Säurebildung im Mund und Funktion des Kaugummikauens in Verbindung bringen; Bewertung von Aussagen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung.
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

II.4. In einer Studie wurden 225 Personen gebeten, drei Monate lang täglich drei Mal ein Kaugummi zu kauen. Die Testpersonen wurden vor und nach dieser Zeit zahnärztlich untersucht. Bei 175 Personen war der Zustand unverändert, bei 40 hatte er sich leicht verbessert, und bei den restlichen 10 geringfügig verschlechtert.
Welche der folgenden Aussagen ist korrekt?

- A Diese Studie beweist, dass Kaugummikauen gut für die Zähne ist.
- B Diese Studie beweist, dass Kaugummikauen keine Wirkung auf die Zähne hat.
- C Diese Studie beweist, dass Kaugummikauen schlecht für die Zähne ist.
- D Diese Studie beweist gar nichts.

Anforderung der Aufgabe

Bewerten: Umgang mit Evidenzen, Bewertung von Aussagen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung.
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

IM CHEMIELABOR

III.1. Der Chemikalienschrank ist gut sortiert. In welcher Aufzählung werden ausschließlich Stoffe genannt, die in Wasser gelöst basisch (alkalisch) reagieren?

- A NaOH, CH₃COOH, H₂SO₄
 B NaOH, CaO, NaHCO₃
 C NaCl, NH₃, KOH
 D CH₃CH₂OH, CaO, HCl

Anforderung der Aufgabe

Wissen: Stoffklasse der Basen;

konvergentes Denken: Stoffe zuordnen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

III.2. (Bei der Zugabe von Natronlauge zu Salzsäure verändert sich der pH-Wert. ...) Kennzeichne, welche der folgenden Kurven die pH-Wert-Änderung richtig wiedergibt!

D (erkennt, dass pH-Wert zu Beginn bei 1 liegt; erkennt, dass pH-Wert über 7 hinaus ansteigt)

Anforderung der Aufgabe

Wissen: Neutralisation, pH-Wert;

Umgang mit grafischen Darstellungen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

III.3. Beim Experimentieren wurde im Abzug Säure verschüttet. Wasser oder Natronlauge? Welchem Ratschlag schließt du dich an? Begründe deine Entscheidung!

Entscheidet sich für Wasser, da man zuvor die Menge Natronlauge genau berechnen müsste, die man benötigt, um die Säure zu neutralisieren, sonst basische, und damit ebenfalls ätzende Reaktion.

Nimmt Wasser, um die Säure stark zu verdünnen (und auf diese Weise den pH-Wert zu erhöhen).

Nimmt Natronlauge, verweist aber darauf, dass der pH-Wert kontrolliert werden muss, damit man nicht zuviel Natronlauge zugibt.

[Ein Verweis auf die freiwerdende Neutralisationswärme als zusätzliche Gefahrenquelle ist nicht notwendig zur vollständigen und korrekten Lösung dieser Aufgabe.]

Anforderung der Aufgabe

Wissen: Neutralisation, Konzentration und Verdünnung;

konvergentes Denken: Problem und Lösungswege zusammenbringen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

III.4. Indikatoren verändern je nach pH-Wert einer Lösung ihre Farbe. Kennzeichne, mit welchem der drei Indikatoren du prüfen kannst, ob Regenwasser pH-neutral ist!

A	Methylorange	Umschlagbereich bei pH 3,1 (rot) – 4,4 (orange)
B	Phenolphthalein	Umschlagbereich bei pH 8,0 (farblos) – 9,8 (violett)
C	Bromthymolblau	Umschlagbereich bei pH 6,2 (gelb) – 7,6 (blau)

Anforderung der Aufgabe

Wissen: Indikatoren, pH-Wert;

konvergentes Denken: Aufgabe und Sachinformation zum pH-Wert in Verbindung bringen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung.
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

III.5. In der linken Tabellenspalte sollen Eigenschaften eingetragen werden, die allen sauren Lösungen gemeinsam sind. In der rechten Tabellenspalte soll je ein Versuch eingetragen werden, mit dem die links genannte Eigenschaft demonstriert werden kann. Ergänze alle Felder der Tabelle!

Eigenschaft/Merkmal einer sauren Lsg.	geeigneter Versuch zur Demonstration
<i>saure Lösungen reagieren mit unedlen Metallen (unter Gas- / H₂-Entwicklung) bzw. saure Lösungen ätzen unedle Metalle Saure Lösungen haben einen pH < 7; saure Lösungen geben mit Indikatoren charakteristische Färbungen; saure Lösungen neutralisieren basische Lösungen; saure Lösungen greifen Kalk an (ätzen); saure Lösungen leiten den elektrischen Strom; ...</i>	<i>Man gibt vorsichtig einige Tropfen der sauren Lösung auf ein unedles Metall (z.B. Zink). pH-Wert ermitteln, wenige Tropfen Indikator zugeben, pH-Wert-Messungen einige Tropfen der sauren Lösung auf Marmor, Kalk, Eierschale, Kreide... geben. Durchführung einer Leitfähigkeitsmessung</i>

Anforderung der Aufgabe

Wissen: Stoffklasse der Säuren, Eigenschaften der Stoffklasse;

konvergentes Denken: Eigenschaften und Nachweise

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung.
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

III.6. Destilliertes Wasser leitet den elektrischen Strom praktisch nicht. Eine saure Lösung hingegen leitet den elektrischen Strom gut. Erläutere bitte das Phänomen mit Hilfe der Teilchen des destillierten Wassers und in sauren Lösung!

Existenz von Ionen in der Lösung als Voraussetzung für elektrische Leitfähigkeit, Erklärung der Leitfähigkeit von sauren Lösungen über das Vorhandensein von H_3O^+ -Ionen und Säure-Anionen, Säuren als Protonendonatoren, Bildung von Protonen und Säure-Anionen, im destillierten Wasser liegen keine Ionen vor, daher keine Leitfähigkeit (expliziter Hinweis auf die geringe Eigendissoziation des Wassers nicht gefordert).

Anforderung der Aufgabe

Wissen: Stoffklasse der Säuren, Protonenbegriff, Voraussetzung für elektrische Leitfähigkeit; konvergentes Denken: Kenntnisse zur elektrischen Leitfähigkeit mit Modellvorstellungen zu sauren Lösungen in Verbindung bringen

Die notwendigen Kenntnisse/Fähigkeiten wurden im Unterricht bereitgestellt:	Ich erwarte, dass der folgende Prozentsatz meiner Schüler/innen diese Aufgabe vollständig und richtig lösen kann:	Diese Art der Aufgabenstellung interessiert folgenden Prozentsatz meiner Schüler/innen:	Die Art der Aufgabe hat in meinem Unterricht eine Bedeutung .
<input type="radio"/> eher ja <input type="radio"/> eher nein	0.....20.....40...50...60.....80.....100	0.....20.....40...50...60.....80.....100	<input type="radio"/> eher ja <input type="radio"/> eher nein

Vielen Dank für Ihre Mitarbeit!

ANHANG ZUM LEHRER-FRAGEBOGEN:

Kurze Zusammenstellung einiger Aufgabenanforderungen

Divergentes Denken

- Es geht um die kreative Produktion alternativer Möglichkeiten, die aber gemeinsam ein Grundverständnis des zu Grunde liegenden Sachverhaltes ausdrücken.

Konvergentes Denken

Wissenselemente logisch miteinander verknüpfen, um zu einer Schlussfolgerung oder Vorhersage zu gelangen

- Es geht um die Fähigkeit, verschiedene Informationsquellen aufeinander beziehen zu können.

Umgang mit grafischen Darstellungen

Änderung des Repräsentationsmodus, Interpretation von grafisch dargebotener Information

- Es geht um die Fähigkeit, relevante Informationen aus grafischen Darstellungen von Daten zu ziehen.

Umgang mit „Sprache“, „sprachlichen Informationen“

Änderung des Repräsentationsmodus, Beschreibung von Beobachtungen, Anwendung von Fach- und Formelsprache

- Es geht darum, Informationen situationsangemessen zu kommunizieren.

Umgang mit Zahlen

- Es geht um die Fähigkeit, (stöchiometrische) Berechnungen anzustellen, Größenverhältnisse abzuschätzen, Informationen aus dargebotenen Zahlen zu ziehen.

Bewertung / Evaluation

Auswahl oder Prioritätensetzung, argumentatives Begründen, Informationen und Quellen bewerten, Folgen abschätzen, Prozess der Urteilsbildung

- Es geht darum, sachliche Informationen in einen Bewertungsprozess einzubinden, Evidenzen zu bewerten.

Die genannten Kompetenzbereiche sind u.a. angelehnt an die sieben Fähigkeitsbereiche, die in der nationalen Ergänzungserhebung zu PISA 2003 dem Bereich „Naturwissenschaften“ zu Grunde lagen – Quelle: Internationales und nationales Rahmenkonzept für die Erfassung von naturwissenschaftlicher Kompetenz in PISA 2003 (Deutsches PISA-Konsortium 2002);
<http://www.ipn.uni-kiel.de/projekte/pisa/NaWiRahmenkonzeptionPISA2003.pdf>

I Ergänzungen zu den Ergebnissen der Hauptstudie

Fragestellung 1.2

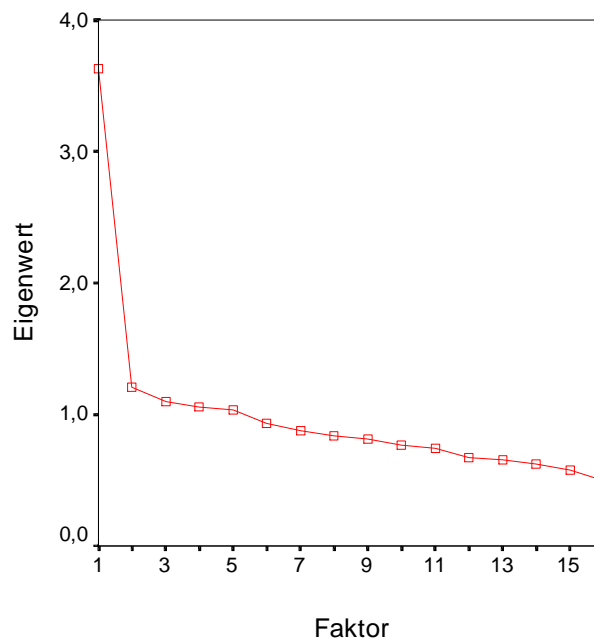


Abb. 19: Screeplot zur Faktorenbestimmung; Darstellung der Eigenwerte.

Aufgabe	Faktor		
	1	2	3
I.2 „Aspekte d. Med.-entwicklung“		.50	
I.3 „Fragen formulieren“			.62
„Internet-Suchbegriffe“		.76	
I.4 „Versuch zum Gasauffangen“	.56		
I.5 „Reaktionsgleichung“	.59		
„Berechnung“	.43		.48
„Informationen einholen“	.37		.48
II.1 „Kurve beschreiben“			.71
„Kurve interpretieren“	.44		
II.2 „Fehlersuche“		.35	
III.1 „basische Stoffe“		.51	
III.2 „pH-Kurven“	.53		
III.3 „Ratschlag“	.59		
III.4 „Indikatoren“	.30	.59	
III.5 „Demonstrationsversuche“	.39	.31	
III.6 „Leitfähigkeit“	.49		

Tab. 51: Ergebnisse der Faktorenanalyse: Darstellung der rotierten Komponentenmatrix mit drei Faktoren unter Berücksichtigung von 16 Items des Tests (interne Konsistenz der Skala: Cronbachs $\alpha = 0,74$); dargestellt sind alle Ladungen $> .30$; Hauptladungen sind durch Fettdruck hervorgehoben; Bestimmung der internen Konsistenz der gefundenen Skalen: $\alpha_{\text{Faktor 1}} = 0,56$, $\alpha_{\text{Faktor 2}} = 0,54$, $\alpha_{\text{Faktor 3}} = 0,49$.

Fragestellung 2.1

Klasse	N	M	SD	Min.	Max.
1	15	6,61	3,045	0,0	10,8
2	17	7,16	3,219	2,5	13,2
3	15	7,99	2,651	4,5	12,5
4	19	5,27	1,585	1,5	8,2
5	24	5,73	3,349	0,0	11,5
6	17	4,36	2,503	0,5	8,8
7	25	6,8	2,835	1,0	11,2
8	23	4,24	2,075	1,0	9,7
9	21	9,4	3,031	4,7	16,0
10	19	6,85	1,265	4,0	8,8
11	27	10,18	1,97	5,2	13,0
12	25	8,86	2,511	4,0	12,8
13	21	9,19	1,965	4,5	12,7
14	24	8,7	1,923	5,5	12,5
15	19	9,18	1,165	7,5	11,7
16	27	11,12	1,356	9,0	14,0
17	21	5,57	1,162	3,2	7,5
	Σ 359	7,63	3,046	0,0	16,0

Tab. 52: Kenndaten der teilnehmenden Klassen zum Abschneiden der einzelnen Klassen im Test (maximal erreichbare Punktzahl = 16): Klassengröße N, Mittelwerte der im Test erreichten Punktzahlen, Standardabweichungen, Minimum und Maximum.

Fragestellung 2.2

Faktor	Cluster				
	1	2	3	4	5
„Schulchemie“	.18	- .65	.70	- .93	.69
„pH-Kurven“	- 1.30	.61	.68	- .04	.28
„Textproduktion“	- .05	.92	- .89	- .56	.54
„modellhafte Vorstellung“	.43	.74	.71	- .51	- .91

Tab. 53a: Clusterzentrenanalyse unter Annahme von fünf Clustern, positive Werte sind durch Fettdruck in der Tabelle hervorgehoben.

Cluster	Probanden
1	75
2	60
3	59
4	80
5	85

Tab. 53b: Anzahl der Fälle in jedem Cluster, N = 359.

K CD-ROM: Unterrichtsmaterialien

Auf der CD befinden sich die Dateien zu den im Rahmen der Arbeit mit den Lehrerinnen und Lehrern der Schulsets erarbeiteten und zusammengestellten Unterrichtsmaterialien.

Lebenslauf

Dorothe Christiansen

Holtener Straße 340 • 24106 Kiel • d.christiansen@web.de

Persönliche Daten:

Geburtsdatum: 5. August 1975
Geburtsort: Eckernförde
Staatsangehörigkeit: deutsch
Familienstand: verheiratet, eine Tochter

Schulbildung:

1982 – 1986 Grundschule in Eckernförde
1986 – 1995 Jungmannschule in Eckernförde

Schulabschluss:

1995 Abitur

Studium:

1995 – 2001 Studium der Fächer Chemie und Deutsch für das Lehramt an
Gymnasien an der Christian-Albrechts-Universität zu Kiel

Abschluss:

27. Juli 2001 Erste Staatsprüfung für die Laufbahn der Studienrätinnen und
Studienräte an Gymnasien

1. Juni 2007

Zweite Staatsprüfung für die Laufbahn der Studienrätinnen und
Studienräte an Gymnasien

Akademische Berufstätigkeit:

August 2001 – Juli 2005

Wissenschaftliche Mitarbeiterin am Leibniz-Institut für die
Pädagogik der Naturwissenschaften (IPN), Abteilung der Didaktik
der Chemie (Arbeitsgebiet: Chemie im Kontext)

August 2005 – Juli 2007

Referendariat am Gymnasium Schloss Plön

seit August 2007

Studienrätin z. A., Gymnasium Schloss Plön

DANKSAGUNG

Diese Arbeit ist nur durch die Unterstützung und Zusprache verschiedener Personen entstanden und fertig gestellt worden, so dass ich an dieser Stelle die Gelegenheit nutze, mich dafür zu bedanken.

Mein Dank gilt in erster Linie Prof. Dr. Ilka Parchmann, die mir die Möglichkeit gegeben hat, im Projekt *Chemie im Kontext* mitzuarbeiten und diese Arbeit zu verfassen, und an die ich mich zu jeder Zeit mit fachwissenschaftlichen Fragen wenden konnte.

Darüber hinaus möchte ich mich bei Prof. Dr. Claudia Nerdel und Dr. Martin Senkbeil bedanken, die mir bei statistischen Fragen weitergeholfen haben.

Für die unermüdliche Fehlersuche im Manuskript bedanke ich mich bei Dr. Iris Stracke und Lars Hollensen.

Danken möchte ich auch der *Chemie-im-Kontext*-Projektgruppe für die Unterstützung während der Arbeit und den Lehrerinnen und Lehrern für die Teilnahme an der Untersuchung. Den Lehrerinnen und Lehrern der Hamburger und rheinland-pfälzischen Schulsets gilt mein besonderer Dank für den anregenden Austausch im Rahmen der Sattreffen. Die Arbeit im Projekt und in den Arbeitsgruppen der Schulsets war nicht immer konfliktfrei, hat aber meinen Blick auf Schule und Unterricht um viele Dimensionen erweitert und mir viel Freude bereitet.

Der Abteilung der Chemiedidaktik des IPN danke ich für die Unterstützung und die aufbauenden Worte während der Erstellung meiner Arbeit. Besonders danken möchte ich Dr. Bianca Steinhoff, die mir gezeigt hat, wie man trotz vielfältiger Anforderungen und Verpflichtungen seinen Weg nicht aus den Augen verliert.

Abschließend möchte ich mich bei meinen Eltern bedanken, die mir den Weg bis hierher überhaupt erst ermöglicht und meine beruflichen Entscheidungen immer unterstützt haben.

Ohne die tatkräftige Unterstützung meiner Eltern und Ingrid Skwarr hätte diese Arbeit nicht ihren Abschluss gefunden. Schließlich danke ich meinem Mann Heiko und meiner Tochter Johanna, die mich immer wieder ermuntert und unterstützt und meinen Blick auf das Wesentliche gelenkt haben.