

OLIVER TEPNER, BURKHARD ROEDER UND INSA MELLE

Effektivität von Aufgaben im Chemieunterricht der Sekundarstufe I

Efficiency of tasks in lower secondary chemistry lessons

Zusammenfassung

In diesem Beitrag wird eine Studie zur Effizienz des Aufgabeneinsatzes im Chemieunterricht der Sekundarstufe I vorgestellt. Dabei wurde im Rahmen eines quasiexperimentellen Pre-/Post-Testdesigns der Frage nachgegangen, ob die Aufgabenverwendung die Lernleistung und die Einstellung der Lernenden im direkten Vergleich zu einer von der gleichen Lehrkraft auf herkömmliche Weise unterrichteten Parallelklasse steigern kann. Während in der Interventionsgruppe Aufgaben selbstständig auf Arbeitsblättern bearbeitet wurden, lernten die Schülerinnen und Schüler der Parallelklasse von ihrer Lehrkraft unterstützt im Plenum. Außerdem wurde nach dem so genannten gekreuzten Parallelklassendesign vorgegangen, um eventuelle Unterschiede zwischen der jeweiligen Interventions- und Kontrollgruppe ausgleichen zu können. Dazu wurden zwei Unterrichtseinheiten nacheinander durchgeführt, sodass jeweils beide Parallelklassen abwechselnd Interventions- und Kontrollgruppe waren. An Vor- und Hauptuntersuchung nahmen 251 bzw. 351 Schülerinnen und Schüler der 9. Klasse an nordrhein-westfälischen Gymnasien teil. Es kann u. a. gezeigt werden, dass der Einsatz von Aufgaben zu einem signifikant höheren Lernerfolg und positiveren Einstellungen führt als herkömmlicher Unterricht.

Schlüsselwörter: Aufgaben, Chemieunterricht, gekreuztes Parallelklassendesign, Effektivität, Einstellung, quantitative Beziehungen bei chemischen Reaktionen

Abstract

This article presents a study about efficiency of task application in chemistry lessons. In order to evaluate students' knowledge and attitudes towards chemistry, tasks, and the regarding teaching unit, a quasi-experimental pre/post test design was used. While the experimental group worked on tasks autonomously using worksheets in a systematic way, the control group learned in plenum supported by their teacher. Additionally, a so-called crossed parallel classes design was used to compensate for eventual differences in confounding variables between experimental and control group. For this purpose, two teaching units were conducted consecutively so that each class was both experimental and control group. The pilot study and the main study were attended by 251 and 351 students, respectively. The study was conducted with ninth-grade grammar-school classes in North Rhine-Westphalia, Germany. Among others, this study reveals significantly better results in the experimental group's knowledge and attitudes compared with those of control group.

Keywords: tasks, chemistry lessons, crossed parallel classes design, efficiency, attitudes, stoichiometry, chemical equations

1 Ausgangssituation

Aufgaben übernehmen nicht erst seit TIMSS und PISA eine wichtige Funktion zur Förderung und Feststellung von Kompetenzen im naturwissenschaftlichen Unterricht (PISA-

Konsortium Deutschland, 2008). Im Chemieunterricht haben Aufgaben insbesondere im Rahmen des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“ begonnen, an Bedeutung zu gewinnen (Bund-Länder-Kommission für

Bildungsplanung und Forschungsförderung [BLK, 1997), was sich beispielsweise in der Entwicklung moderner und vielfältiger Aufgabenformate äußert (Stäudel, 2004a). Bisher fehlt allerdings die empirische Basis für die Effizienz des Einsatzes von Aufgaben im Chemieunterricht. In diesem Artikel werden daher die Ergebnisse einer Studie vorgestellt, welche die Ziele hatte, einerseits Aufgaben zu entwickeln und andererseits ihre Effektivität zu untersuchen. Im Rahmen einer Vergleichsgruppenuntersuchung wurde der Unterricht mit verstärktem Aufgabeneinsatz im Vergleich zu herkömmlich unterrichteten Parallelklassen evaluiert. Dabei meint der Begriff „herkömmlich“ den Unterricht, welchen Lehrende im Regelfall und ohne weitere Vorgaben durchführen. Nähere Informationen bezüglich des Aufgabeneinsatzes finden sich in Kapitel 4.1.

2 Forschungsstand

2.1 Effektivität von Aufgaben im Unterricht

Im Folgenden wird ein Überblick über die empirische Forschung zum Einsatz von Aufgaben im Unterricht gegeben. Während der Aufgabenbegriff in den zitierten Arbeiten nicht immer einheitlich definiert ist, wird er im Rahmen der vorliegenden Studie später präzisiert und beispielsweise um den Aspekt des Schriftlichen ergänzt (Kapitel 4).

In den vorliegenden Studien wurden Aufgaben im Rahmen von Hausaufgaben und in Kombination mit Arbeitsblättern erforscht, bei denen der Aspekt der Selbstständigkeit von herausragender Bedeutung ist. Allerdings sind die berichteten Effekte in den bisherigen Vergleichsstudien zur Hausaufgabenthematik entweder nicht signifikant oder nicht eindeutig auf die Intervention zurückzuführen (Bonham, Deardorff & Beichner, 2003; Cooper, Robinson & Patall, 2006; Hascher & Bischof, 2000). Eine Ausnahme stellt hier die Arbeit von Nicolai (2005) dar, welche die lernförderliche Wirkung von experimentellen Hausaufgaben,

insbesondere für schwächere Schülerinnen und Schüler, am Beispiel des Chemieunterrichts belegt. Ein wesentlicher Faktor bei dem Bearbeiten von (Haus-)aufgaben scheint die Anstrengungsbereitschaft der Schülerinnen und Schüler zu sein, die mit dem Lernerfolg positiv korreliert, während dieser mit der reinen Bearbeitungszeit negativ korreliert (Trautwein & Köller, 2003; Trautwein, Lüdtke, Schnyder & Niggli, 2006). Dies steht im Gegensatz zu Forschungsergebnissen, nach denen die Hausaufgabenbearbeitungszeit mit dem Lernerfolg direkt korreliert, wie sie z. B. Cooper, Robinson und Patall (2006) berichten.

Eine weitere Möglichkeit, Aufgaben weitgehend selbstständig zu bearbeiten und für den Lernprozess fruchtbar zu machen, stellt das Lernen aus Lösungsbeispielen dar, zu dem sich ebenfalls widersprüchliche Effekte finden. Beispielsweise erweisen sich Lösungsbeispiele in Kombination mit Lernprogrammen insbesondere dann als effektiv, wenn die Vor- und Nachteile der vorgestellten Beispiele selbst erklärt werden, also eine hohe Eigenaktivität der Lernenden, in diesem Fall 80 Lehramtsstudierende, erfordern (Schworm & Renkl, 2002). Im Gegensatz dazu können Catrambone und Yuasa (2006) allerdings keinen signifikant höheren Lernzuwachs nachweisen, wenn Lösungsbeispiele durch eigene Zusammenfassungen bzw. nähere Ausführungen aktiv ergänzt werden, obwohl die den einzelnen Aufgaben gewidmete Zeit höher war als in der passiv lernenden Kontrollgruppe, der die Ergänzungen vorgegeben wurden. Möglicherweise hängt dies von der Struktur der Aufgaben ab und lässt sich nicht verallgemeinern. Die Relevanz des eigenständigen Arbeitens kann in einem weiteren Teil der Studie jedoch in Ansätzen bestätigt werden: Transferaufgaben erledigten die aktiven Lerner signifikant schneller als die passiv Lernenden (ebd.). Der Aspekt der (unterstützten) Selbstständigkeit ist auch in aktuellen Studien zu instruktionalen Erklärungen zentral, in denen die aktive (mentale) Verarbeitung der Lern-

prozesse von entscheidender Bedeutung ist (Berthold & Renkl, 2008). Zudem kann nach Mackensen-Friedrichs (2005) die Effektivität von Beispielaufgaben u. a. durch eine Erhöhung des Übungsanteils weiter gesteigert werden.

Auch für die Effektivität des Einsatzes von Aufgaben auf Arbeitsblättern im Sinne von Paper-Pencil-Aufgaben existieren nur wenige empirische Belege. So kann beispielsweise Kobayashi (2006) in einer Metastudie den positiven Effekt von selbstverfassten Anmerkungen zu einem Text nachweisen. Insgesamt liegen keine einheitlichen Untersuchungsergebnisse zum Einsatz von Aufgaben im Unterricht vor; allerdings zeigen die zitierten Studien Ansätze auf, unter denen Aufgaben Lernerfolg steigernd eingesetzt werden könnten. Ein lernerfolgsrelevanter Faktor, der sich aus allen Studien ableiten lässt, scheint insbesondere die Eigenaktivität der Schülerinnen und Schüler zu sein, welche in den seit vielen Jahren bekannten Vorschlägen zur Unterrichtsgestaltung wie z. B. aktives Anwenden und Stillarbeit zum Ausdruck kommen (Rosenshine, 1995).

2.2 Aufgaben im Fach Chemie

Während sich die bisher vorgestellten Ergebnisse primär auf die allgemein pädagogische Ebene beziehen, sollen im Folgenden chemie- oder naturwissenschaftsdidaktische Befunde im Mittelpunkt stehen.

Im Unterschied beispielsweise zum Mathematik-, Sprach- oder Kunstunterricht existiert im Fach Chemie keine Aufgabentradition, welche primär die schriftliche Bearbeitung meint (Stäudel, 2003). Noch vor sechs Jahren stellen Nentwig, Christiansen und Steinhoff pointierend fest: „Speziell für den Chemieunterricht fehlt bisher eine systematische Auseinandersetzung mit dem Themengebiet ‚Aufgaben‘, wie man sie zum Beispiel für den Physikunterricht findet“ (Nentwig, Christiansen & Steinhoff, 2004, S. 22). Erst in jüngerer Zeit lassen sich vermehrt Publikationen für die Chemie bzw.

die Naturwissenschaften allgemein finden, welche u. a. als Reaktion auf die internationalen Schulleistungstests PISA und TIMSS zu verstehen sind und sich beispielsweise auf das DFG-Projekt „Aufgaben im naturwissenschaftlichen Unterricht“ beziehen (Tiemann, Rumann, Jatzwauk & Sandmann, 2006). Im Rahmen dieses Projekts haben Tiemann, Draxler und Labusch (2004) u. a. untersucht, inwieweit Aufgaben Unterricht beeinflussen können. Zudem wurde erhoben, welche Funktionen Aufgaben im Unterricht einnehmen (Tiemann, Rumann, Jatzwauk & Schabram, 2006).

Das in den letzten Jahren gestiegene Forschungsinteresse an Aufgaben bedeutet nicht, dass es für den Chemieunterricht bis dato gar keine Aufgaben gegeben hat (vgl. auch Stäudel, 2004b). Stäudel stellt in einem kurzen historischen Aufriss Bezüge zu Aufgaben her, die von 1940 bis heute in Schulbüchern zu finden sind. Stäudel zufolge liegen Veränderungen weniger in modifizierten Aufgabenformaten als in der unterrichtsmethodischen Einordnung: Aktuell dienen Aufgaben weniger der Verständnis- und Leistungsüberprüfung als dem eigentlichen Lernen anhand eigener Erfahrungen. Interessant ist in diesem Zusammenhang, dass gerade das Format der PISA-Aufgaben originär zur Kompetenzmessung konzipiert worden ist und nun häufig im Schulalltag zur Erarbeitung neuen Wissens und zum Lernen verwendet wird (ebd.). Verbunden mit diesem Perspektivenwechsel ist eine teilweise Abkehr von der Wissenskonsolidierung zugunsten der Wissenserarbeitung unter dem Schlagwort „Neue Aufgabenkultur“.

Zur Auswirkung des Einsatzes von Aufgaben im Hinblick auf den Lernerfolg existieren vereinzelte Studien mit uneinheitlichen Ergebnissen. Beispielsweise können Slunt und Giancarlo in ihrer Studie keine eindeutig positiven Effekte von Aufgaben zur Vorbereitung auf universitäre Chemie-Vorlesungen berichten (Slunt & Giancarlo, 2004). Bereits einige Jahre zuvor untersuchten Pöpping und Melle (2001) den Stellenwert von Aufgaben, welche als Übungen verwendet

werden. Sie fanden u. a. heraus, dass nur in jeder sechsten der beobachteten 130 Chemiestunden geübt wird. Auch die Tatsache, dass über 80 Prozent der Übungen primär Wiederholungen sind, verdeutlicht den geringen Stellenwert, den Übungen bzw. entsprechend intendierte Aufgaben allgemein im Chemieunterricht haben. Wie eine Befragung von Tepner und Melle ergab, sind eher die Lehrkräfte der Meinung, es werde ausreichend im Chemieunterricht geübt, welche die mündlichen Wiederholungen am Stundenanfang zu den Übungen zählen, obwohl diese kaum Übungscharakter aufweisen (Tepner & Melle, 2004). Oftmals sind bei der Wiederholung der Inhalte der letzten Stunde nur wenige Schülerinnen und Schüler aktiv, während der Großteil einer Klasse bestenfalls rezipiert, sodass ein effektives Üben hier fraglich erscheint (ebd.). Möglicherweise liegt eine Überschätzung des tatsächlichen Übungsanteils durch Lehrkräfte an einer uneinheitlichen Auffassung des Übungsbegriffs.

Neben dem zunehmenden Forschungsinteresse zu Aufgaben im Chemieunterricht finden sich auch eine Reihe von eher unterrichtspraktisch angelegten Überlegungen (z. B. Nentwig et al., 2004; Woest, 2004). Für den Physikunterricht fordern von Aufschnaiter & von Aufschnaiter (2001), dass Aufgaben ein integraler Bestandteil von Unterricht sein und mehr als 50 Prozent der Unterrichtszeit ausmachen sollten. Dabei gehen die Autoren von einem eher allgemeinen Aufgabenbegriff aus, der eine hohe Eigenaktivität voraussetzt und Lernende Konzeptualisierungen selbst entwickeln lässt. Entsprechend dieser Auffassung wäre es wünschenswert, wenn Aufgaben auch im Chemieunterricht eine höhere Bedeutung erlangten.

3 Ziele und Hypothesen

Wesentliche Ziele der Untersuchung sind einerseits der grundsätzliche Nachweis der Effektivität des Einsatzes von Aufgaben im

Chemieunterricht und andererseits das Präzisieren der Bedingungen, unter denen dies der Fall sein kann. Dabei wird unter Effektivität der erreichte Lernzuwachs der Schülerinnen und Schüler zu einer bestimmten Unterrichtseinheit im Vergleich zu einer Kontrollgruppe verstanden. Zudem sollen Auswirkungen auf motivationaler Ebene seitens der Lernenden erhoben werden, um der anzunehmenden wechselseitigen Beeinflussung von Lernerfolg und Motivation Rechnung zu tragen (Kapitel 4). Im Rahmen der Studie sollen folgende Hypothesen geprüft werden:

1. Lernende, welche im Chemieunterricht selbstständig und systematisch Aufgaben bearbeiten, erzielen einen höheren Lernzuwachs als Lernende einer herkömmlich unterrichteten Kontrollgruppe.
- 2a. Lernende, welche im Chemieunterricht selbstständig und systematisch Aufgaben bearbeiten, bewerten den Chemieunterricht allgemein anders als die Kontrollgruppe ohne selbstständig bearbeitete Aufgaben.
- 2b. Lernende, welche im Chemieunterricht selbstständig und systematisch Aufgaben bearbeiten, bewerten die durchgeführte Unterrichtsreihe anders als die Kontrollgruppe ohne selbstständig bearbeitete Aufgaben.
- 2c. Lernende, welche im Chemieunterricht selbstständig und systematisch Aufgaben bearbeiten, bewerten den Einsatz von Aufgaben anders als die Kontrollgruppe ohne selbstständig bearbeitete Aufgaben.

Darüber hinaus wurden folgende Nebenypothesen generiert:

1. Das selbstständige und systematische Bearbeiten von Aufgaben im Chemieunterricht fördert die Leistung von Jungen und Mädchen gleichermaßen.
2. Das selbstständige und systematische Bearbeiten von Aufgaben im Chemieunterricht fördert die Leistung von leistungs-

starken und leistungsschwachen Schülerinnen und Schüler gleichermaßen.

Vor- und Hauptuntersuchung sind grundsätzlich identisch, sodass aus Platzgründen im Folgenden nur auf die Hauptuntersuchung eingegangen wird. Die zentralen Begriffe „selbstständig“ und „systematisch“ werden im Folgenden erläutert.

4 Einsatz und Konzeption der Aufgaben

4.1 Aufgabeneinsatz

Zur Überprüfung der Effektivität des Aufgabeneinsatzes sollten die Interventions- und Kontrollgruppe (IG und KG) unter verschiedenen Bedingungen unterrichtet werden. In beiden Gruppen sollten jedoch keine Experimente in die Unterrichtsreihe eingebunden werden, um einen zusätzlichen Einfluss auf die Einstellungsänderungen aufgrund ihrer motivierenden Eigenschaften zu vermeiden und evtl. Änderungen allein auf die Intervention zurückzuführen.

Unter der Prämisse, der Einsatz von Aufgaben sei unter bestimmten Bedingungen besonders effektiv, wird in dieser Studie unter dem Begriff Aufgabe die schriftliche Auseinandersetzung mit einer ebenfalls schriftlich formulierten Arbeitsaufforderung (Bruder, 2006; Tiemann et al., 2006) verstanden. Die vorrangig in relativ kurzer Zeit zu bearbeitenden Aufgaben (Häußler & Lind, 1998) wurden ähnlich dem Forschungsansatz von Hascher und Bischof (2000) in die reguläre Unterrichtszeit integriert und nur in Ausnahmefällen als Hausaufgaben vergeben, um in der Stunde begonnene Aufgaben zu vollenden. Damit sollten in nahezu jeder Unterrichtsstunde Aufgaben sowohl zur Erarbeitung als auch zur Anwendung und Übung eingesetzt werden.

Um dem Aspekt des selbstständigen und eigenverantwortlichen Lernens Rechnung zu tragen, wurden die verwendeten Aufgaben vorrangig schriftlich in Einzelarbeit

und – falls eine Schülerin oder ein Schüler nicht mehr alleine zurechtkam – auch in Partnerarbeit bearbeitet. Damit konnten die Lernenden ihr Lerntempo individuell bestimmen und – z.B. für den Fall, dass nicht mehr genügend Zeit zur Verfügung steht oder bestimmte Aufgaben zu schwierig sind – eine Auswahl der angebotenen Aufgaben treffen (Bleichroth, 1998). Die Aufgaben wurden allen Schülerinnen und Schülern grundsätzlich auf Arbeitsblättern zur Verfügung gestellt und im Rahmen einer Unterrichtsreihe zum Thema „Quantitative Beziehungen bei chemischen Reaktionen“ eingesetzt. Letzteres bestand aus den Teilthemen *Molbegriff*, *Molare Masse*, *Aufstellen von Reaktionsgleichungen* und *Rechnen in der Chemie*. Das Thema wurde ausgewählt, da Chemielehrende in einer vorausgegangenen Befragung den Wunsch nach mehr anspruchsvollen Aufgaben bzw. Übungen insbesondere zu diesem theorieelastigen Thema äußerten (Tepner & Melle, 2003). Die Annahme, ein systematisches Bearbeiten von Aufgaben führt insbesondere in diesem Themenfeld zu einem hohen Lernerfolg, können Fach, Endres und Parchmann (2006) bestätigen. Diese untersuchten das Thema „Quantitative Betrachtungen chemischer Reaktionen“ bzw. die Lösungsstrategien von 17 Schülerinnen und Schülern beim stöchiometrischen Rechnen. Demnach erzielten vorrangig die Lernenden richtige Lösungen, welche algorithmisch vorgehen. Zudem gibt es Hinweise, dass gestufte Hilfen für chemische Fragestellungen hilfreich sein können (Fach, de Boer & Parchmann, 2007).

Anlass für den Einsatz von Aufgaben in der Form klassischer Arbeitsblätter war die Intention, ein unkompliziertes und flexibel einsetzbares Medium zu verwenden und möglichst geringe Eingriffe in den Schulalltag vorzunehmen und somit die Aufgabenverwendung später ggf. dauerhaft in der Schule implementieren zu können (Ballstedt, 1993; Uihlein, Graf & Klee, 2003). Im Gegensatz zu oftmals geliehenen Schulbüchern bieten Arbeitsblätter – insbesondere

Chemisches Rechnen	Infoblatt										
<p>Idee: Aus chemischen Formeln und Reaktionsgleichungen lassen sich die <i>Stoffmengenverhältnisse</i> der einzelnen Elemente ableiten. Z. B. reagieren Aluminium und Sauerstoff im Verhältnis 2:3 zu Aluminiumoxid (Al₂O₃). Das heißt: 1 mol Al₂O₃ besteht aus 2 mol Al-Atomen und 3 mol O-Atomen. Diese Informationen lassen sich auch zur Berechnung der beteiligten Massen nutzen.</p>											
<p>Dazu benötigst du wieder die Formel $M = \frac{m}{n}$</p> <div style="display: flex; align-items: center; justify-content: center; margin-left: 100px;"> <div style="text-align: center; margin-right: 10px;"> \swarrow Masse </div> <div style="text-align: center; margin-right: 10px;"> \nwarrow Stoffmenge </div> </div> <p>und du musst wissen, wie man die molare Masse <i>M</i> einer Verbindung (z. B. Al₂O₃) ermittelt.*</p>											
<p>Beispielaufgabe: Aluminium wird nicht nur in der Automobilindustrie, sondern auch für die Herstellung von Aluminiumfolien benötigt. Wie viel kg Aluminium (Al) lassen sich aus 100 kg Aluminiumoxid (Al₂O₃) gewinnen?</p>											
<p>Lösungsschritte:</p>	<div style="border: 1px solid gray; padding: 5px; margin-bottom: 10px; background-color: #f0f0f0;"> <p>Angaben: $M(O) = 16 \text{ g/mol}$ $M(Al) = 27 \text{ g/mol}$</p> </div> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; padding: 5px;">1. Größen mit Einheiten zusammentragen:</td> <td style="padding: 5px;"> <u>gesucht:</u> $m(Al) = ? \text{ kg}$ <u>bekannt:</u> $m(Al_2O_3) = 100 \text{ kg}$ </td> </tr> <tr> <td style="padding: 5px;">2. Molare Massen der betreffenden Substanzen ermitteln:</td> <td style="padding: 5px;"> $M(Al) = 27 \text{ g/mol}$ $M(Al_2O_3) = (2 \cdot 27 + 3 \cdot 16) \text{ g/mol} = 102 \text{ g/mol}$ </td> </tr> <tr> <td style="padding: 5px;">3. Stoffmenge <i>n</i> der bekannten Substanz (Al₂O₃) über die molare Masse <i>M</i> ermitteln:</td> <td style="padding: 5px;"> $n = m/M = 100000 \text{ g} / (102 \text{ g/mol}) = 980,4 \text{ mol}$ </td> </tr> <tr> <td style="padding: 5px;">4. Stoffmenge der gesuchten Substanz (Al) bestimmen: <small>(Die Stoffmenge <i>n</i>(Al) ist doppelt so groß wie die von Al₂O₃, da pro Formeleinheit Al₂O₃ 2 Teilchen Al gewonnen werden.)</small></td> <td style="padding: 5px;"> $n(Al) = 2 \cdot n(Al_2O_3) = 1960,8 \text{ mol}$ </td> </tr> <tr> <td style="padding: 5px;">5. Masse der gesuchten Substanz über <i>M</i> ermitteln:</td> <td style="padding: 5px;"> $m(Al) = n(Al) \cdot M(Al)$ $= 1960,8 \text{ mol} \cdot 27 \text{ g/mol}$ $= 52942 \text{ g} = \underline{52,942 \text{ kg}}$ </td> </tr> </table>	1. Größen mit Einheiten zusammentragen:	<u>gesucht:</u> $m(Al) = ? \text{ kg}$ <u>bekannt:</u> $m(Al_2O_3) = 100 \text{ kg}$	2. Molare Massen der betreffenden Substanzen ermitteln:	$M(Al) = 27 \text{ g/mol}$ $M(Al_2O_3) = (2 \cdot 27 + 3 \cdot 16) \text{ g/mol} = 102 \text{ g/mol}$	3. Stoffmenge <i>n</i> der bekannten Substanz (Al ₂ O ₃) über die molare Masse <i>M</i> ermitteln:	$n = m/M = 100000 \text{ g} / (102 \text{ g/mol}) = 980,4 \text{ mol}$	4. Stoffmenge der gesuchten Substanz (Al) bestimmen: <small>(Die Stoffmenge <i>n</i>(Al) ist doppelt so groß wie die von Al₂O₃, da pro Formeleinheit Al₂O₃ 2 Teilchen Al gewonnen werden.)</small>	$n(Al) = 2 \cdot n(Al_2O_3) = 1960,8 \text{ mol}$	5. Masse der gesuchten Substanz über <i>M</i> ermitteln:	$m(Al) = n(Al) \cdot M(Al)$ $= 1960,8 \text{ mol} \cdot 27 \text{ g/mol}$ $= 52942 \text{ g} = \underline{52,942 \text{ kg}}$
1. Größen mit Einheiten zusammentragen:	<u>gesucht:</u> $m(Al) = ? \text{ kg}$ <u>bekannt:</u> $m(Al_2O_3) = 100 \text{ kg}$										
2. Molare Massen der betreffenden Substanzen ermitteln:	$M(Al) = 27 \text{ g/mol}$ $M(Al_2O_3) = (2 \cdot 27 + 3 \cdot 16) \text{ g/mol} = 102 \text{ g/mol}$										
3. Stoffmenge <i>n</i> der bekannten Substanz (Al ₂ O ₃) über die molare Masse <i>M</i> ermitteln:	$n = m/M = 100000 \text{ g} / (102 \text{ g/mol}) = 980,4 \text{ mol}$										
4. Stoffmenge der gesuchten Substanz (Al) bestimmen: <small>(Die Stoffmenge <i>n</i>(Al) ist doppelt so groß wie die von Al₂O₃, da pro Formeleinheit Al₂O₃ 2 Teilchen Al gewonnen werden.)</small>	$n(Al) = 2 \cdot n(Al_2O_3) = 1960,8 \text{ mol}$										
5. Masse der gesuchten Substanz über <i>M</i> ermitteln:	$m(Al) = n(Al) \cdot M(Al)$ $= 1960,8 \text{ mol} \cdot 27 \text{ g/mol}$ $= 52942 \text{ g} = \underline{52,942 \text{ kg}}$										

Abb. 2: Beispiel eines Informationsblattes.

Die verwendeten Infoblätter (Abbildung 2) ergänzten die Erläuterungen der Lehrkraft und legten teilweise die Voraussetzungen zur selbstständigen Bearbeitung der nachfolgenden Aufgaben. So konnten die Schülerinnen und Schüler ein Schema zum Aufstellen von Reaktionsgleichungen oder das Berechnen z. B. von Massen und Stoffmengen mithilfe von Reaktionsgleichungen nachvollziehen. Die Infoblätter durften im Plenum auch in den Kontrollgruppen eingesetzt werden.

Um einen möglichst realistischen Vergleich mit der bestehenden Situation im Chemieunterricht zu erhalten, sollte die Kontrollgruppe von der gleichen Lehrkraft auf „herkömmliche“ Weise unterrichtet werden.

„Herkömmlich“ meinte nicht nur den weit verbreiteten mehr oder weniger lehrerzentrierten Frontalunterricht, sondern allgemein auch den Unterrichtsstil, den die jeweilige Lehrperson auch ohne Teilnahme an dem Projekt angewendet hätte. Dabei durften und sollten eigene Aufgaben verwendet werden, soweit sie ohne Beteiligung an der Studie auch eingesetzt worden wären, wenn auch nicht ausschließlich in Einzelarbeit auf Arbeitsblättern, sondern z. B. in einem Unterrichtsgespräch unter Nutzung z. B. der Tafel. Vor der Untersuchung wurde mit den Lehrkräften besprochen, in welcher Weise diese gewöhnlich das Thema unterrichten. Bei allen Unterschieden, die zwischen dem Unterricht verschiedener Lehrkräfte immer

vorhanden sind, teilten sie uns mit, dass sie normalerweise Unterricht im Plenum durchführten, Aufgaben aber, wenn überhaupt, primär an der Tafel besprächen. Diese würden ggf. eher als Hausaufgabe denn im Unterricht eingesetzt. Der exakte Umfang der Unterrichtsreihe wurde von den Lehrkräften individuell bestimmt, allerdings sollten IG und KG jeweils gleich lang und inhaltsgleich sein. Die zeitliche und inhaltliche Vergleichbarkeit von IG und KG wurde anhand von Unterrichtsprotokollationsbögen und Audioaufnahmen sichergestellt. Die beschriebene Kontrollgruppenbedingung diente dem Vergleich mit einer herkömmlichen Unterrichtssituation und wird im Folgenden als *real* bezeichnet. Um Rückschlüsse auf den Einfluss der Selbstständigkeit bei der Aufgabenbearbeitung zu ziehen zu können und auszuschließen, dass Effekte allein auf die in den Aufgaben behandelten Inhalte zurückzuführen sein könnten, existierte eine weitere Kontrollgruppenbedingung, welche im Folgenden als *angeleitet* bezeichnet wird. In dieser bearbeitete die KG die gleichen Aufgaben wie die IG, allerdings nicht in Einzelarbeit, sondern im Plenum unter An-

leitung der Lehrkraft. Die Musterlösungen und Infoblätter wurden mit den Lehrerinnen und Lehrern an der Tafel besprochen. In Tabelle 1 findet sich eine Übersicht über die Vorgaben der Untersuchung.

4.2 Gestaltungskriterien der Arbeitsblätter

Im Hinblick auf eine effektive und praktikable Auswertung der Studie sind sowohl die Testaufgaben als auch die im Unterricht eingesetzten Aufgaben als Basis der Testvorbereitung weitgehend geschlossen konzipiert und lassen sich mit herkömmlichen Aufgabentypen vergleichen, wie sie aus der Mathematik bekannt sind. Anhand eines Arbeitsblattbeispiels sollen weitere Gestaltungskriterien erläutert werden.

Im Sinne der Strukturierung und Orientierung dient das einheitliche Layout der Kopfzeile dazu, Lernenden auf den ersten Blick kenntlich zu machen, worum es geht und ermöglicht die Zuordnung mehrerer Arbeitsblätter zu einem Thema (Eckert, 1980). Über die Angabe des Namens und Datums

Tab. 1: Vorgaben für den Unterricht in der Interventions- und Kontrollgruppe

Versuchsbedingung	Interventionsgruppe	Kontrollgruppe
1 „selbstständig vs. real“	<ul style="list-style-type: none"> • Schaffen der Voraussetzungen zum selbstständigen Lösen der Aufgaben • Bearbeitung möglichst aller bereit gestellten Aufgaben • in Einzel- bzw. Partnerarbeit • schriftlich auf dem Arbeitsblatt • Bearbeitung nur im Ausnahmefall als Hausaufgabe • Lösungsblätter als Hilfe bzw. Feedback 	<ul style="list-style-type: none"> • reale Unterrichtssituation: herkömmlicher Unterricht und Aufgabeneinsatz • eigene Aufgaben in gewöhnlich von der Lehrkraft gewähltem Umfang zulässig; einzige Einschränkung: nicht ausschließlich in Einzelarbeit
2 „selbstständig vs. angeleitet“	s. o.	<ul style="list-style-type: none"> • gleiche Aufgaben wie IG • Bearbeitung mit der Lehrkraft an der Tafel bzw. im Plenum • ohne Lösungsblätter

erfolgt zudem eine indirekte Identifikation mit den Materialien, und es soll eine gewisse Verbindlichkeit erzeugt werden, das betreffende Arbeitsblatt verantwortungsvoll auszufüllen.

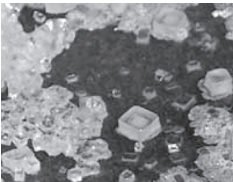
Um Interesse für die Bearbeitung einer Aufgabe zu wecken, werden Alltagsbezüge hergestellt, von denen eine motivierende Wirkung auf Schülerinnen und Schüler angenommen wird (Fach, Kandt & Parchmann, 2006; Fechner, 2009; Fischer & Draxler, 2001).

Mit einer möglichst kurzen und prägnanten Formulierung wird dem Wunsch von Lernenden und Lehrenden nach einer kurzen

und klaren Aufgabenstellung entsprochen (Fach et al., 2006), sodass eine eigenständige Bearbeitung ohne evtl. störende Rückfragen gewährleistet sein sollte. Im Regelfall enthält die Aufgabenstellung nur einen Operator wie z. B. „Beschreibe“, „Begründe“, „Berechne“ (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK], 2004). Nur in Formulierungen, welche ohnehin eindeutig zu verstehen sind, wird aus stilistischen Gründen auf einen Operator verzichtet und es werden – wie in dem Beispiel – direkte Fragen gestellt (Abbildung 3).

Datum:	Chemisches Rechnen	Vorname:
Klasse:		Erster Buchstabe des Nachnamens:

Aufgabe 1
Natriumchlorid („Salz“, NaCl) wird beispielsweise als Steinsalz „unter Tage“ abgebaut. Im Labor jedoch kann Natriumchlorid direkt aus den Elementen (Na) und Chlor (Cl₂) hergestellt werden. Die Reaktionsgleichung lautet: $2 \text{Na} + \text{Cl}_2 \rightarrow 2 \text{NaCl}$. Wie viel mol NaCl können maximal aus 6 mol Natrium (Na) und 2 mol Chlor (Cl₂) entstehen?




Natriumchlorid-Kristalle
(eigene Aufnahme)

Aufgabe 2
Silicium (Si) findet bei der Herstellung von Solarzellen und Computern Anwendung. Es wird aus Quarz (SiO₂) hergestellt. Wie viel Gramm Quarz werden benötigt, um 10 g Silicium herzustellen? Um die Aufgabe zu lösen, kannst du auf einem ähnlichen Wege vorgehen, wie du ihn in dem „Aluminiumoxid-Beispiel“ kennen gelernt hast.

Angaben:

$M(\text{O}) = 16 \text{ g/mol}$

$M(\text{Si}) = 28 \text{ g/mol}$



Schwimmbaddach mit Solarzellen
www.solarserver.de/solarmagazin/index.html, Zugriff am 12.1.2004

einheitliche Kopfzeile

lebensnaher/ aktueller Kontext

eindeutige Aufgabenstellung

überwiegend mit Fotos gestaltet

Raum für Bearbeitung auf dem gleichen Blatt

ansteigender Schwierigkeitsgrad

Anwenden deklarativen u. prozeduralen Wissens

einheitliche Umrandung

Abb. 3: Arbeitsblattbeispiel.

	Klasse 1	Klasse 2	Evaluationsinstrumente
1. Abschnitt (HU1)	Pre-Tests	Pre-Tests	Fähigkeitstest (S) Einstellungstests (S) Wissenstest (S)
	Aufgaben 1. Themenblock	Kontrollbedingungen	Audioaufzeichnung (L) bzw. Stundenprotokoll (L)
	Post-Tests	Post-Tests	Einstellungstests (S) Wissenstest (S)
2. Abschnitt (HU2)	Pre-Tests	Pre-Tests	Einstellungstests (S) Wissenstest (S)
	Kontrollbedingungen	Aufgaben 2. Themenblock	Audioaufzeichnung (L) bzw. Stundenprotokoll (L)
	Post-Tests	Post-Tests	Einstellungstests (S) Wissenstest (S)

Abb. 4: Untersuchungsdesign (gekreuztes Parallelklassendesign), □ = IG, ■ = KG, S = Schüler, L = Lehrer.

Der ansteigende Schwierigkeitsgrad wurde im Wesentlichen an der Komplexität der Aufgaben festgemacht. Z. B. mussten beim Aufstellen von Reaktionsgleichungen in den ersten Aufgaben entweder nur Koeffizienten oder Indizes angepasst werden. Erst anschließend war das vollständige Aufstellen der gesamten Gleichung erforderlich.

5 Design und Stichprobe

Die Studie wurde im so genannten „gekreuzten Parallelklassendesign“ konzipiert, welches als Kombination zweier quasiexperimenteller Kontrollgruppenuntersuchungen aufzufassen ist. Während Aufgaben in einer Klasse selbstständig bearbeitet wurden (Interventionsgruppe, „IG“, in Abbildung 4 hellgrau schattiert), diente die Parallelklasse als Kontrollgruppe („KG“, dunkelgrau schattiert), in welcher kein entsprechender Einsatz von Aufgaben stattfand. Die zwei Parallelklassen wurden jeweils von derselben Lehrkraft unterrichtet, somit konnte die Störvariable „unterschiedliche Lehrkraft“ eliminiert werden. Nach der Durchführung des ersten Untersuchungsabschnitts (HU1) wurden Interven-

tions- und Kontrollgruppe getauscht. Da beide Klassen einmal IG und ein anderes Mal KG waren, konnten mögliche Unterschiede aufgrund personenbezogener Störvariablen in beiden Untersuchungsbedingungen annähernd gleich ausgeprägt verteilt werden (Bortz, 2005).

Dem Ansatz liegt folgende Überlegung zugrunde: Wenn eine Klasse bereits im Vorfeld der Untersuchung deutlich besser wäre als ihre Parallelklasse, dann sollte sie sowohl als IG als auch als KG im Wissenstest erfolgreicher sein. Dies wäre zu erwarten, wenn der Effekt aufgrund der Klassenunterschiede größer wäre als der Treatment-Effekt, also der Effekt, den der Aufgabeneinsatz zur Folge hat. Falls jedoch die jeweilige IG in beiden Abschnitten (HU1 und HU2) erfolgreicher wäre, sollte die Intervention einen bedeutsameren Effekt haben als eventuelle Klassenunterschiede. Die Hauptuntersuchung lässt sich also in zwei formal identische Abschnitte unterteilen, deren Struktur in Abbildung 4 dargestellt ist. Die Dauer der Unterrichtseinheiten betrug im Regelfall ca. sechs Unterrichtsstunden, IG und KG wurden jeweils mit der identischen Stundenanzahl unterrichtet.

Mittels zweier weiterer Maßnahmen wurden eventuelle Unterschiede zwischen den Klassen nivelliert bzw. konservativ getestet: Einerseits wird das individuelle Vorwissen über die Berechnung von Residuen berücksichtigt, andererseits bildete grundsätzlich diejenige Parallelklasse die erste Interventionsgruppe, welche beim Fähigkeitstest die jeweils schlechtere war. Mit dieser konservativen, d. h. gegen die Forschungshypothese gerichteten, Anlage der Studie konnte sichergestellt werden, dass mögliche Effekte zugunsten der Hypothese nicht besseren Voraussetzungen der ersten Interventionsgruppe geschuldet waren. Zudem ist die HU1 im Gegensatz zur HU2 nicht von der vorherigen Untersuchung beeinflusst. Damit ist insbesondere die HU1 methodisch einwandfrei konzipiert, sodass auf dieser der Fokus der Auswertung und Interpretation liegt, während die bei der Auswertung

der HU2 resultierenden Ergebnisse nachfolgend mit der HU1 in Bezug gesetzt werden. Die Wissenstests und der Einstellungstest zum Fach Chemie wurden im Pre-/Post-Test-Design verwendet. Falls es zwischen den beiden Untersuchungsabschnitten eine Pause gegeben hat – die Entscheidung blieb den Lehrkräften überlassen – wurde auch der Einstellungstest pre vor dem zweiten Abschnitt durchgeführt, andernfalls nur der thematisch neue Wissenstest pre. Ergänzend wurde zu Beginn ein Fähigkeitstest durchgeführt, welcher der Kontrolle der kognitiven Fähigkeiten der Klassen diene. Stundenprotokolle und Audioaufzeichnungen dienen der Überprüfung der inhaltlichen Validität der Tests und der Vergleichbarkeit von IG und KG. Eine Übersicht über die Daten der Stichprobe und der verwendeten Aufgaben findet sich in Tabelle 2 und Tabelle 3.

Tab. 2: Daten der Stichproben, m=männlich, w=weiblich

	Voruntersuchung	Hauptuntersuchung
Bundesland	Nordrhein-Westfalen	
Schulform	Gymnasium	
Jahrgangsstufe	9	
Anzahl der Klassen	9	12
Anzahl der Lehrkräfte	5 (w: 2 / m: 3)	6 (w: 2 / m: 4)
Anzahl der Schülerinnen und Schüler	249 (w: 130 / m: 119)	322 (w: 193 / m: 129)
Anzahl der Aufgaben (IG)	33	35
Anzahl der Beispielaufgaben	3	-
Anzahl der Infoblätter	1	3

Tab. 3: Übersicht über die in der HU verwendeten Aufgaben

	Thema	Aufgabenzahl HU
1. Abschnitt	<i>Molbegriff</i>	6 (+ 1 Infoblatt)
	<i>Molare Masse</i>	13
	gesamt	19 (+ 1 Infoblatt)
2. Abschnitt	<i>Aufstellen von Reaktionsgleichungen</i>	8 (+ 1 Infoblatt)
	<i>Rechnen in der Chemie</i>	8 (+ 1 Infoblatt)
	gesamt	16 (+ 2 Infoblätter)
Gesamt		35 (+ 3 Infoblätter)

6 Erhebungsinstrumente

6.1 Wissenstests

Die Wissenstests wurden im Multiple-Choice-Format konzipiert und umfassten 19 Aufgaben zu den Teilthemen *Molbegriff* und *Molare Masse* in der HU1 bzw. 18 Aufgaben zu den Teilthemen *Aufstellen von Reaktionsgleichungen* und *Rechnen in der Chemie* in der HU2. Neben deklarativem wurde vorrangig prozedurales Wissen (z. B. Aufstellen von Reaktionsgleichungen, Berechnen von Stoffmengen) geprüft. Um eine Antwortalternative – beispielsweise zum Aufstellen von Reaktionsgleichungen – als richtig oder falsch einordnen zu können, reicht es nicht aus, auswendig gelerntes Fachwissen zu reproduzieren, sondern dieses in Form von Prozessen zu verknüpfen und anzuwenden. Die Entscheidung basiert also auf vorangegangenen Prozessen (Paris, Lipson & Wixson, 1983). Die Tests wurden als Pre- und Post-Test eingesetzt, wobei der Post-Test zensurenrelevant war. Beim Post-Test wurde nach der häuslichen Lernzeit für diesen Test gefragt. Die Tests waren in ca. 20 Minuten von den Schülerinnen und Schülern zu bearbeiten. Eine Beispielaufgabe aus dem Wissenstest findet sich in Abbildung 5.

Pro Item existierten eine richtige Lösung und vier Distraktoren. Das Maß für die individuelle Leistung im Wissenstest war der Anteil

der richtig gelösten Items. Dieser Score liegt somit zwischen 0 und 1. Um eventuell unterschiedliche Niveaus im Vorwissen bei der Interpretation der Post-Testergebnisse zu berücksichtigen, wurden zur Auswertung der Daten z-standardisierte Residuen berechnet. Diese verstehen sich als Abweichung des empirisch ermittelten Post-Testwerts vom vorhergesagten Wert, welcher auf Basis der linearen Regression errechnet wurde und Ausdruck der Stichproben bezogenen Abhängigkeit des Post-Testwertes vom Pre-Testwert ist (Bortz, 2005).

6.2 Einstellungstests

Die entwickelten Tests maßen die allgemeine Einstellung zum Fach Chemie (pre/post), die Einschätzung der durchgeführten Unterrichtsreihe (post) und die Einschätzung des Aufgabeneinsatzes (post). Für den Test zum Fach Chemie seien beispielhaft folgende Items aufgeführt: „Im Chemieunterricht fühle ich mich überfordert“, „Im Chemieunterricht lerne ich Dinge, die ich später gebrauchen kann“ und „Ich gehe ungern zum Chemieunterricht“. Die Unterrichtsreihe wurde z. B. anhand folgender Items eingeschätzt: „Ich fand den Unterricht der letzten 6 Stunden angenehm“, „Ich habe das Gefühl, in den letzten 6 Stunden wenig gelernt zu haben“ und „Der Unterricht der letzten 6 Stunden hat mir Spaß gemacht“.

8) Ammoniak (NH_3) reagiert mit Sauerstoff (O_2) zu Stickstoff (N_2) und Wasser. Gib die Reaktionsgleichung an!

- $2 \text{NH}_3 + 3 \text{O}_2 \rightarrow \text{N}_2 + 6 \text{H}_2\text{O}$
- $4 \text{NH}_3 + 3 \text{O}_2 \rightarrow 2 \text{N}_2 + 6 \text{H}_2\text{O}$
- $4 \text{NH}_3 + 6 \text{O}_2 \rightarrow 2 \text{N}_2 + 12 \text{H}_2\text{O}$
- $2 \text{NH}_3 + \text{O}_2 \rightarrow 2 \text{N}_2 + \text{H}_2\text{O}$
- $2 \text{NH}_3 + 6 \text{O}_2 \rightarrow \text{N}_2 + 6 \text{H}_2\text{O}$

Abb. 5: Beispielaufgabe Wissenstest Post, 2. Untersuchungsabschnitt.

Der Test zum Aufgabeneinsatz bestand aus der Instruktion, dass sich das Wort „Aufgaben“ auf die in den letzten Unterrichtsstunden kennen gelernten Aufgaben und nicht auf Experimente bezieht, und umfasste u. a. folgende Items: „Die gestellten Aufgaben fand ich interessant“, „Ich fand die verwendeten Aufgaben überflüssig“ und „Bei den Aufgaben konnte ich endlich zeigen, was ich in Chemie wirklich kann“. Bei der Konzeption wurde sich primär auf Roeder (2006), Bühner (2004) und Nolen (2003) bezogen. Die Tests bestanden aus 25 bis 28 geschlossenen Items (formuliert als Statements), welche auf einer 5-stufigen Likert-Skala von „sehr zutreffend“ bis „sehr unzutreffend“ zu beantworten waren.

Zur Auswertung der Daten wurden zunächst die Items, welche negativ und gegen die Dimensionsrichtung formuliert worden waren, umkodiert, indem sie um die Skalenmitte gespiegelt wurden (Brosius, 2004). Anschließend wurde zu jedem Item der exakte Median berechnet (Clauß, Finze & Partzsch, 2004). Anschließend wurde eine Dichotomisierung vorgenommen, bei der ein Item mit 0 kodiert wurde, wenn es eine Punktwertung oberhalb des Medians erhalten hatte, andernfalls mit 1. Somit repräsentieren die erhaltenen Scores den Anteil der Antworten, welche den Items zustimmen. Höhere Werte drücken also eine positivere Einstellung aus. Gegenüber der mittelwertbasierten hat die mediandichotomisierte Auswertung Vorteile: Einerseits werden die Ordinalskalierung und damit die unterschiedlichen Abstände der Likertskala angemessen berücksichtigt, da eine Mittelwertberechnung streng betrachtet kardinalskalierte Daten erfordert (Bortz, 2005; Clauß et al., 2004). Extreme Verteilungen aufgrund von sehr eindeutigen Antworten werden besser berücksichtigt, indem jedes Item anhand des Medians (also anhand einer Rangreihe) für oder gegen die Dimensionsrichtung gewertet wird und wie die übrigen zur Varianz der gesamten Skala beiträgt. So werden auch Items mit zunächst geringer und großer Streuung gleich gewichtet. Zudem umgeht das Verfahren

der Dichotomisierung das Problem, Probandinnen und Probanden über die Vorgabe einer geradzahigen Antwortskala zu einer eindeutigen Entscheidung zwischen Zustimmung und Ablehnung zu zwingen. Da die Referenz für die Dichotomisierung a posteriori anhand der bestehenden Datenbasis festgelegt wird und auch für ungeradzahige Skalen errechnet werden kann, werden nicht eindeutige Antworten wie „mittel“ als positiv oder negativ gewertet.

Den beschriebenen Vorteilen der Dichotomisierung steht allerdings gegenüber, dass Informationen aufgrund der Reduktion in 0 und 1 verloren werden (Bühner, 2004). Vor dem Hintergrund einer eindeutigeren Separation zwischen positiv und negativ eingestellten Schülerinnen und Schülern, der besseren Vergleichbarkeit mit anderen Tests über die errechneten Scores und der selbstverständlich identischen Anwendung auf IG und KG erscheint dieser Nachteil akzeptabel.

6.3 Fähigkeitstest

Der verwendete Fähigkeitstest CFT 20 („Culture Fair Test“) diente primär der Feststellung eventueller Unterschiede bezüglich Interventions- und Kontrollgruppe (Weiss, 1998). Es wurde sich auf den ersten Testteil beschränkt, sodass die Testdauer auf ca. 20 Minuten reduziert werden konnte.

6.4 Stundenprotokolle und Audioaufzeichnungen

Die Protokollationsbögen und Audioaufnahmen hatten zum Ziel, unerwartete Untersuchungsergebnisse im Nachhinein aufklären zu können. Die Lehrenden wurden gebeten, entweder nach jeder Unterrichtsstunde einen Fragebogen zum Stundeninhalt auszufüllen oder die eigene Stimme mit Hilfe eines um den Hals getragenen mp3-Recorders aufzuzeichnen, um über die Klassenbucheinträge hinausgehende Informationen zu gene-

rieren. Dadurch konnte die grundsätzliche Vergleichbarkeit hinsichtlich der Lernzeit und der behandelten Inhalte der Parallelklassen überprüft werden. Die Auswertung der Audioaufnahmen von drei Lehrkräften ergab, dass ein Lehrer IG und KG nicht entsprechend der Vorgaben unterrichtet hat. Als Konsequenz wurden die betreffenden Daten aus der Wertung genommen.

7 Ergebnisse der 1. Hauptuntersuchung

7.1 Wirksamkeit des Aufgabeneinsatzes

Im Folgenden werden primär die Ergebnisse der 1. Hauptuntersuchung (HU1) präsentiert, die eine in sich abgeschlossene Untersuchung der aufgestellten Hypothesen erlaubt. Die zweite Hauptuntersuchung (HU2) ist durch die vorherige beeinflusst und wird daher separat ausgewertet und interpretiert. Aus Platzgründen werden die Ergebnisse der Voruntersuchung nicht berichtet, sie werden durch die der Hauptuntersuchung bestätigt. In der Auswertung werden die Kontrollbedingungen *real* und *angeleitet* mit der Interventionsbedingung verglichen. Die beiden Interventionsgruppen werden analog den Kontrollgruppen als IG_{real} und IG_{angeleitet} bezeichnet, um deutlich zu machen, dass es sich um zwei verschiedene Stichproben – allerdings mit identischen Vorgaben – handelt.

Obwohl einige Teilstichproben nicht absolut normalverteilt und varianzhomogen sind, können parametrische Testverfahren wie die Varianzanalyse aufgrund der Stichprobengrößen angewendet werden (Bortz, 2005; Box, 1954). Zur Absicherung wurden die referierten Ergebnisse mit nonparametrischen Verfahren wie dem Mann-Whitney-U- und Wilcoxon- anstelle der t-Tests überprüft und bestätigt (Tepner, 2008).

In der HU1 kann die erste Hypothese, dass systematisch in den Unterricht integrierte Aufgaben einen höheren Lernzuwachs bewirken, bestätigt werden. Unabhängig von den Kontrollgruppenbedingungen (*real/angeleitet*) hat die Interventionsgruppe signifikant besser abgeschnitten als die jeweilige Kontrollgruppe (Stichprobe *real*: $t(107) = 3.368$; $p = .001$; $d = 0.61$ und Stichprobe *angeleitet*: $t(144) = 2.808$; $p = .003$; $d = 0.46$). Die angegebenen Werte des Signifikanztests beziehen sich auf die Residuen. Da die Interventionsgruppe auch unter der angeleiteten Bedingung besser als die Kontrollgruppe ist, scheint der „bloße“ Einsatz von Aufgaben nicht besonders effektiv zu sein. Dann müsste die KG_{angeleitet} ähnlich erfolgreich abgeschnitten haben wie die IG_{angeleitet}, weil in beiden Teilstichproben identische Aufgaben eingesetzt worden sind. Dies kann als Evidenz für die Wirkung des selbstständigen Bearbeitens von Aufgaben gewertet werden. In Tabelle 4 und Abbildung 6 sind die deskriptiven Er-

Tab. 4: Ergebnisse Wissenstests der HU1. Grundsätzlich wurden alle Werte für die Stichproben *real* und *angeleitet* separat ermittelt. Die Stichprobengröße bezieht sich jeweils auf die Residuenberechnung. Cronbach's Alpha wurde für den Post-Test bestimmt. n = Stichprobenumfang; M = Mittelwert; SD = Standardabweichung

Stichprobe	Gruppe	n	Pre-Test		Post-Test		Residuen		Cronbach's Alpha
			M	SD	M	SD	M	SD	
<i>real</i>	IG	53	.371	0.122	.809	0.149	.315	0.801	.75
	KG	58	.354	0.154	.699	0.182	-.288	1.023	
<i>angeleitet</i>	IG	70	.356	0.173	.769	0.174	.236	0.950	.77
	KG	76	.381	0.144	.695	0.191	-.218	1.043	

gebnisse und das Säulendiagramm der HU1 aufgeführt.

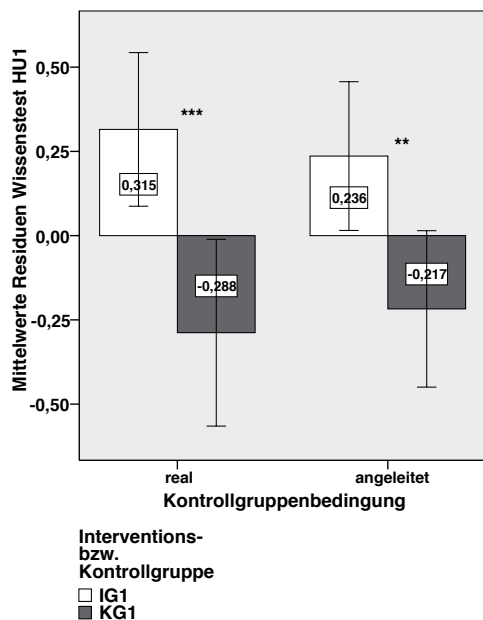


Abb. 6: Säulendiagramm Ergebnisse Wissenstests der HU1. Die dargestellten Mittelwerte der Residuen wurden für die Stichproben *real* und *angeleitet* separat berechnet. Fehlerbalken kennzeichnen das 95%-Konfidenzintervall, Sternchen signifikante Unterschiede.

7.2 Interaktionseffekte

Mit der Analyse der Wechselwirkungen von Aufgabeneinsatz und Vornote in Bezug auf die erreichten residualen Lernzuwächse, sollte der Frage nachgegangen werden, ob die Aufgabenbearbeitung sowohl für leistungsstärkere als auch leistungsschwächere Schülerinnen und Schüler gleichermaßen geeignet ist. Dazu wurden die Vorzensuren im Fach Chemie anhand des Medians ($Mdn=2.76$) dichotomisiert (Clauß et al., 2004): Zu den leistungsstärkeren Schülerinnen und Schülern der Stichprobe wurden die mit 1 und 2 Vorbenoteten gezählt, zu den leistungsschwächeren die mit 3 und 4. Die Noten 5 und 6 waren nicht vergeben worden. Auch wenn Zensuren nicht in jedem Fall als

geeignetes Maß für Leistung erachtet werden können, erscheinen die Vorzensuren für die Einteilung in leistungsstärker und -schwächer ausreichend aussagekräftig.

Mit Hilfe der 2-faktoriellen Varianzanalyse lassen sich in der HU1 keine signifikanten Wechselwirkungen zwischen der Zeugnisnote in Chemie und der Gruppenzugehörigkeit nachweisen (Stichprobe *real*: $F<1$ und Stichprobe *angeleitet* $F<1$).

In Bezug auf eventuelle Wechselwirkungen von Treatment und Geschlecht der Probanden sind ebenfalls keine signifikanten Interaktionseffekte nachweisbar (Stichprobe *real*: $F<1$ und Stichprobe *angeleitet*: $F(1,142)=1.219$; $p=.271$; partielles $\eta^2=.009$). Damit profitieren Mädchen und Jungen vom Aufgabeneinsatz gleichermaßen, sodass auch diese entsprechende Ausgangshypothese bestätigt wird.

Da alle eingesetzten Materialien auch zur Vorbereitung auf den notenrelevanten Test verwendet werden durften, wäre es denkbar, dass das Testergebnis nicht nur vom Lernerfolg während der Unterrichtsstunden, sondern auch von der häuslichen Vorbereitung, gemessen an der angegebenen Lernzeit für den Wissenstest, abhängig ist. Allerdings lassen sich keine Wechselwirkungen zwischen der Lernzeit für den Test und dem Treatment nachweisen (Stichprobe *real*: $F<1$ und Stichprobe *angeleitet* $F<1$).

7.3 Einstellung zum Fach Chemie

Zur Erfassung der Unterschiede zwischen Interventions- und Kontrollgruppe in Bezug auf die allgemeine Einstellung zum Fach Chemie wurden analog den Wissenstests anhand der Pre- und Posttests Residuen berechnet. Es finden sich keine signifikanten Unterschiede zwischen der Interventions- und Kontrollgruppe (Stichprobe *real*: $t(168)=-1.496$; $p=.136$; $d=0.23$ und der Stichprobe *angeleitet*: $t(143)=0.583$; $p=.561$; $d=0.10$). Auch unabhängig vom Treatment finden sich keine Einstellungsänderungen. Innerhalb der einzelnen Teilstichproben

Tab. 5: Wirkung des Aufgabeneinsatzes auf die Einstellung zum Fach Chemie der HU1. Grundsätzlich wurden alle Werte für die Stichproben *real* und *angeleitet* separat ermittelt. Die Stichprobengröße bezieht sich jeweils auf die Residuenberechnung. Cronbach's Alpha wurde für den Post-Test bestimmt

Stichprobe	Gruppe	n	Pre-Test		Post-Test		Residuen		Cronbach's Alpha
			M	SD	M	SD	M	SD	
<i>real</i>	IG	83	.511	0.266	.510	0.257	-.117	0.994	.93
	KG	87	.572	0.204	.596	0.232	.111	0.993	
<i>angeleitet</i>	IG	70	.543	0.249	.574	0.234	.050	1.004	.94
	KG	75	.563	0.248	.573	0.264	-.047	0.994	

Tab. 6: Wirkung des Aufgabeneinsatzes auf die Einschätzung der durchgeführten Unterrichtsreihe der HU1. Grundsätzlich wurden alle Werte für die Stichproben *real* und *angeleitet* separat ermittelt

Stichprobe	Gruppe	n	Post-Test		Cronbach's Alpha
			M	SD	
<i>real</i>	IG	83	.483	0.288	.94
	KG	87	.563	0.263	
<i>angeleitet</i>	IG	70	.693	0.221	.94
	KG	75	.583	0.290	

IG_{real} , $IG_{angeleitet}$, KG_{real} und $KG_{angeleitet}$ sind keine signifikanten Unterschiede zwischen Pre- und Posttest nachweisbar (IG_{real} : $t < 1$; KG_{real} : $t(86) = -1.443$; $p = .153$; $d = 0.07$; $IG_{angeleitet}$: $t(69) = -1.601$; $p = .114$; $d = 0.09$; $KG_{angeleitet}$: $t < 1$).

Damit muss die eingangs aufgestellte Hypothese, die Einstellung zum Chemieunterricht könne durch den Einsatz von Aufgaben beeinflusst werden, negiert werden. Die deskriptiven Daten sind in Tabelle 5 aufgeführt.

7.4 Einstellung zur durchgeführten Unterrichtsreihe

In Tabelle 6 sind die Ergebnisse des Einstellungstests zu der zuvor durchgeführten Unterrichtsreihe dargestellt. Diese gestalten sich uneinheitlich. Während in der Stichprobe *angeleitet* signifikante Unterschiede zwischen IG und KG zugunsten der IG zu

finden sind ($t(138) = 2.591$; $p = .011$; $d = 0.42$), kehrt sich dieses Bild in der Stichprobe *real* beinahe um. Dort scheint die KG die Unterrichtsreihe etwas besser zu beurteilen, auch wenn die Differenz zur IG nicht signifikant ist ($t(169) = -1.901$; $p = .059$; $d = 0.29$). Dieses widersprüchliche Ergebnis resultiert im Wesentlichen aus Unterschieden zwischen den Interventionsgruppen *real* und *angeleitet*. Da sich beide Stichproben nur im Treatment der Kontrollgruppe unterscheiden, könnten die Differenzen zwischen beiden IGs nur mit Lehrer- bzw. Schülereffekten sinnvoll erklärt werden.

7.5 Einstellung zum Aufgabeneinsatz

Die Darstellung der Ergebnisse zum Aufgabeneinsatz erfolgt analog den zuvor präsentierten Einstellungstests (Tabelle 7). Der Aufgabeneinsatz wird in den Stichpro-

Tab. 7: Wirkung des Aufgabeneinsatzes auf die Einschätzung der verwendeten Aufgaben der HU1. Grundsätzlich wurden alle Werte für die Stichproben *real* und *angeleitet* separat ermittelt. Die Stichprobengröße und Cronbach's Alpha beziehen sich jeweils auf den Post-Test

Stichprobe	Gruppe	n	Post-Test		Cronbach's Alpha
			M	SD	
<i>real</i>	IG	83	.570	0.265	.92
	KG	87	.545	0.255	
<i>angeleitet</i>	IG	70	.623	0.228	.90
	KG	75	.562	0.244	

ben *real* und *angeleitet* ähnlich, allerdings nicht signifikant positiver bewertet als von der KG (Stichprobe *real*: $t(168)=0.626$; $p=.532$; $d=0.10$ und Stichprobe *angeleitet*: $t(143)=1.559$; $p=.121$; $d=0.26$). Möglicherweise ist die Interventionsdauer zu gering, um eine Änderung der Meinung auf dieser Ebene hervorzurufen. Ein weiterer Grund könnte sein, dass die in der IG und KG verwendeten Aufgaben einander trotz des unterschiedlichen Einsatzes zu sehr ähneln, um den Einsatz der Aufgaben in IG und KG unterschiedlich einzuschätzen.

7.6 Zusammenfassung der HU1

Die Ergebnisse der ersten Hauptuntersuchung können wie folgt zusammengefasst werden: Während sich die Effektivität des Aufgabeneinsatzes eindeutig nachweisen lässt, scheint das Treatment weder auf der Einstellungsebene eine allgemeine Änderung der Einstellung zum Fach Chemie zu bewirken, noch werden die Unterrichtsreihe und der Aufgabeneinsatz einheitlich positiver beurteilt. Mit der Durchführung der zweiten Hauptuntersuchung im Anschluss können Hinweise auf die Stabilität der ermittelten Einstellungen und Lernerfolge gewonnen werden. Im folgenden Kapitel werden daher die Ergebnisse der HU2 auch mit denen der HU1 in Beziehung gesetzt.

8 Ergebnisse der 2. Hauptuntersuchung

8.1 Wirksamkeit des Aufgabeneinsatzes

Im Gegensatz zur HU1 ergibt die Analyse der Stichproben *real* und *angeleitet* für die HU2 keine signifikanten Unterschiede bezüglich der Residuen zwischen IG und KG (Stichprobe *real*: $t<1$; Stichprobe *angeleitet*: $t<1$, Tabelle 8). Damit lässt sich die aufgestellte Hypothese, der selbstständige und systematische Einsatz von Aufgaben führe zu einem höheren Lernerfolg, für die HU2 zunächst nicht bestätigen. Ein möglicher Erklärungsansatz kann in allgemeinen Kompetenzen liegen, welche von den Schülerinnen und Schülern während der HU1 erworben worden sind, die aber im betreffenden Pre-Test noch nicht zum Tragen kommen, da gewisse Detailkenntnisse fehlen, die erst im 2. Abschnitt erworben werden.

Diese Kompetenzen können beispielsweise die Analyse des Zusammenhangs zwischen Aufgabenstellung und -beantwortung oder allgemein das selbstständige Arbeiten sein, welches zwar nicht unbedingt während des Treatments, aber für die Aufgabenbearbeitung im Rahmen des Tests relevant ist. Zudem hat die zweite KG eine größere Wissensbasis erworben, auf der nun neues Wissen vernetzt werden kann. Weiter unten folgen ausführlichere Erläuterungen. Zur Analyse der HU2 wurde daher eine Anova mit dem Faktor Treatment (IG/KG) zur Vorhersage des residualen Lernzuwachses un-

ter Kontrolle des in der HU1 erworbenen Vorwissens berechnet. Als Kovariaten gingen der Pre-Test der HU2 sowie der Post-Test der HU1 ein. Während der Unterschied im residualen Lernzuwachs der Stichprobe *real* nicht signifikant ist ($F(1,100)=0.979$; $p=.163$; partielles $\eta^2=.010$), finden sich

signifikante Differenzen in der Stichprobe *angeleitet* ($F(1,138)=3.800$; $p=.027$; partielles $\eta^2=.027$).

Die in Abbildung 7 dargestellten Unterschiede zwischen den Residuen von $IG_{\text{angeleitet}}$ und $KG_{\text{angeleitet}}$ sind signifikant und können als Indizien gewertet werden, dass die HU2 vom Ergebnis der HU1 beeinflusst wird; die Post-Testergebnisse aus der HU1 wurden bei der Residuenberechnung berücksichtigt.

Dieser Erklärungsansatz wird dadurch gestützt, dass der Lernerfolg in den Stichproben *real* und *angeleitet* der HU1 signifikant mit dem Erfolg der HU2 korreliert (Stichprobe *real*: $r=.262$; $p=.008$ und Stichprobe *angeleitet*: $r=.435$; $p=.000$). Schülerinnen und Schüler, welche bei der HU1 erfolgreich waren, schneiden auch während der HU2 als KG_2 eher gut ab.

Deuten lässt sich dies, indem berücksichtigt wird, dass die KG_2 gegenüber der IG_2 einen Kompetenzvorsprung aufweist, der für die IG_2 nur schwer einzuholen ist. Zudem hat die KG_2 einen Erfahrungsvorsprung in Bezug auf das selbstständige Arbeiten gegenüber der IG_2 und für die HU2 nutzbare Kompetenzen erworben. Auch lässt möglicherweise eine allgemeine Motivationsabnahme im Laufe der Untersuchung die Interventionsmaßnahme während der HU2 generell weniger effektiv sein als zuvor.

Zusammenfassend kann davon ausgegangen werden, dass sich der während der HU1 nachgewiesene Vorteil der IG gegenüber

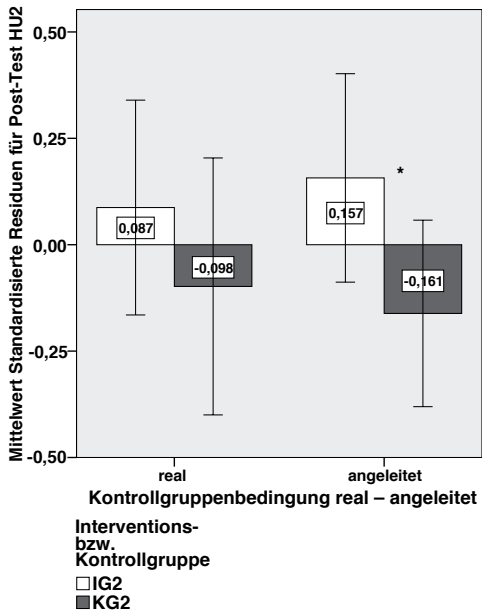


Abb. 7: Säulendiagramm Ergebnisse Wissenstest der HU2. Die dargestellten Mittelwerte der Residuen wurden aus den Post-Testergebnissen der HU1 bzw. HU2 und für die Stichproben *real* und *angeleitet* separat berechnet. Fehlerbalken kennzeichnen das 95%-Konfidenzintervall, Sternchen signifikante Unterschiede.

Tab. 8: Ergebnisse Wissenstests der HU2. Grundsätzlich wurden alle Werte für die Stichproben *real* und *angeleitet* separat ermittelt. Die Stichprobengröße bezieht sich jeweils auf die Residuenberechnung. Cronbach's Alpha wurde für den Post-Test bestimmt

Stichprobe	Gruppe	n	Pre-Test		Post-Test		Residuen		Cronbach's Alpha
			M	SD	M	SD	M	SD	
<i>real</i>	IG	55	.219	0.114	.648	0.185	.021	0.834	.68
	KG	50	.205	0.098	.635	0.195	-.023	0.893	
<i>angeleitet</i>	IG	72	.210	0.089	.543	0.263	.047	1.192	.82
	KG	70	.203	0.085	.531	0.204	-.048	0.889	

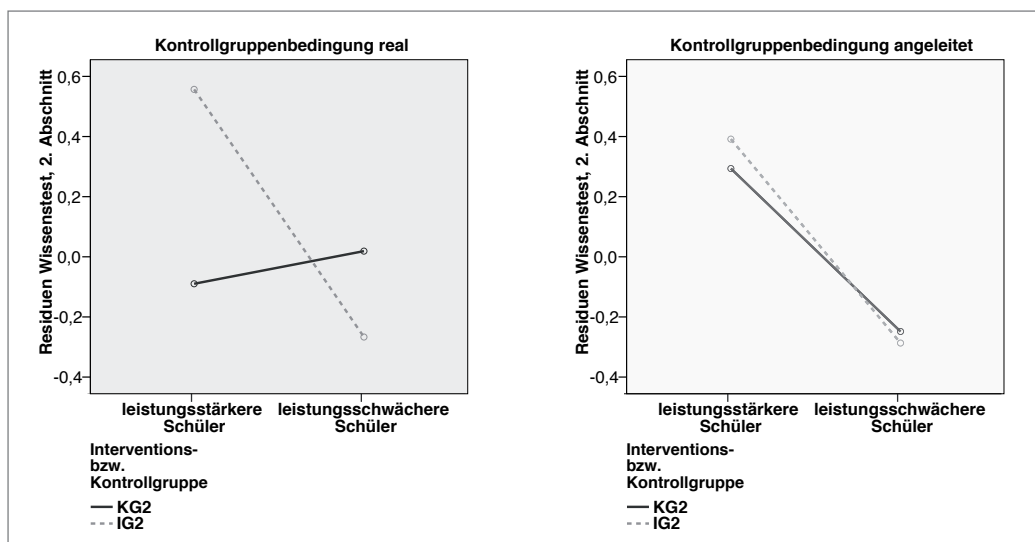


Abb. 8: Wechselwirkungen Vornote/Treatment der HU2; links: Stichprobe real; rechts: Stichprobe angeleitet. Grundsätzlich wurden alle Werte für die Stichproben *real* und *angeleitet* separat ermittelt.

ber der KG nicht nur auf das relevante Wissen bezieht und sich auf die HU2 günstig auswirkt, sondern erworbene grundlegende Kompetenzen in Bezug auf eine systematische Aufgabenbearbeitung positive Effekte auch bei einem neuen Thema zeigen. Der Einsatz der Aufgaben ist bezüglich des Wissenszuwachses in beiden Abschnitten effektiv, wenn das Vorwissen aus dem ersten Abschnitt berücksichtigt wird; Dieser ist nur in der Stichprobe *real* nicht signifikant (Abbildung 7). Ursächlich können nicht alleine die Aufgaben sein, sondern die Selbstständigkeit wirkt sich positiv aus, wie das Design *angeleitet* zeigt.

8.2 Interaktionseffekte

Im Gegensatz zur HU1 gestaltet sich die Analyse der Wechselwirkungen des Treatments mit den Variablen Vornote, Geschlecht und Lernzeit unterschiedlich. Während sich in der Stichprobe *angeleitet* keine signifikanten Interaktionseffekte von Chemienote und Treatment finden ($F < 1$), kehrt sich dieser Befund für die Stichprobe, deren Kontrollgruppe unter realen Bedingungen unter-

richtet wurde, um ($F(1,87) = 4.231$; $p = .043$; partielles $\eta^2 = .046$), Abbildung 8.

In der IG_{2real} profitieren die leistungsstärkeren Lernenden vom für sie neuen Aufgabeneinsatz, während die leistungsschwächeren dieses nicht können (Abbildung 8). Die neue KG_{2real} hingegen ermöglicht als ehemalige IG_{1real} tendenziell leistungsschwächeren Schülertypen, relativ erfolgreich zu sein. Möglicherweise haben die leistungsschwächeren Schülerinnen und Schüler der IG_{2real} als KG_{1real} keine Chance, mit den theoretisch anspruchsvolleren Themen *Aufstellen von Reaktionsgleichungen* und *Rechnen in der Chemie* selbstständig klar zu kommen, insbesondere, wenn sie zuvor nicht gelernt haben, anhand von Aufgaben eigenverantwortlich zu lernen. Allgemein ist zu erwarten, dass Schülerinnen und Schüler nur dann selbstständig lernen können, wenn die zu lernenden Inhalte nicht zu sehr von dem entfernt sind, was sie bereits können. Während der HU2 hat der Kontrollgruppenunterricht ohne Aufgaben möglicherweise einen nivellierenden Effekt. Die Lehrkraft kann zwar die leistungsstärkeren Schülerinnen und Schüler nicht so stark fördern, die leistungsschwä-

chere bekommen aber mehr Hilfe. Außerdem dürfte die KG_{2real} als frühere IG_{1real} vom vorherigen Aufgabeneinsatz profitiert haben. Welcher Effekt allerdings stärker zu gewichten ist, kann nicht exakt bestimmt werden. Bei theoretisch sehr anspruchsvollen Themen kann die Vorbereitung und Hilfe durch den Lehrenden insbesondere für weniger leistungsstarke Lernende relevant sein.

Die Unabhängigkeit der Treatmentwirkung von Geschlecht und Lernzeit für den Test wird in allen Teilstichproben bestätigt (jeweils $F < 1$).

8.3 Einstellung zum Fach Chemie

In Tabelle 9 sind die Ergebnisse der HU2 dargestellt. Analog der HU1 sind die Unterschiede zwischen der Interventions- und Kontrollgruppe nicht signifikant (Stichprobe *real*: $t(155) = -1.463$; $p = .146$; $d = 0.23$ und Stichprobe *angeleitet*: $t < 1$). Damit werden die Befunde der HU1 bestätigt.

Allerdings scheint die Unterrichtseinheit unabhängig vom Treatment einen grundsätzlichen Einfluss auf die Einschätzung zum Fach Chemie zu haben. Während sich nach Durchführung der HU1 keine Veränderung der Einstellung beobachten lässt, zeigt sich (Tabelle 9), dass die Scores von Vor- und Nachtest in den vier Teilstichproben der HU2 signifikant unterschiedlich sind (IG_{real} : $t(80) = 3.956$; $p = .000$; $d = 0.21$; KG_{real} :

$T(75) = 4.062$; $p = .000$; $d = 0.27$; $IG_{angeleitet}$: $t(62) = 4.879$; $p = .000$; $d = 0.24$; $KG_{angeleitet}$: $t(69) = 3.948$; $p = .000$; $d = 0.29$).

Das Ergebnis deckt sich mit vergleichbaren Studien, in denen nach relativ kurzen Interventionszeiten keine grundlegenden Einstellungsänderungen zum Fach zu erwarten sind (z.B. Tepner, 2007), Kapitel 7.3. In diesem Fall scheint sich jedoch die sehr lange Beschäftigung der Schülerinnen und Schüler mit dem theorieelastigen Thema nach Beendigung der HU2 in IG und KG beider Stichproben gleichermaßen negativ auszuwirken. Dieser Befund zeigt sich auch bei den anderen zwei Einstellungstests der Hauptuntersuchung.

8.4 Einschätzung der durchgeführten Unterrichtsreihe

Bei der Einschätzung der durchgeführten Unterrichtsreihe zeigt sich ein uneinheitliches Bild (Tabelle 10). Während sich in der Stichprobe *real* signifikante Unterschiede zwischen IG und KG finden ($t(155) = -3.022$; $p = .003$; $d = 0.47$), ist dieses in der Stichprobe *angeleitet* nicht der Fall ($t < 1$). Die Ergebnisse lassen sich möglicherweise mit den in der Kontrollgruppe *angeleitet* verwendeten Aufgaben deuten, indem diese per se einen positiven Einfluss auf die Einschätzung der Unterrichtsreihe haben, sodass der Unterschied zwischen IG und KG geringer ausfällt als in der Stichprobe *real*.

Tab. 9: Wirkung des Aufgabeneinsatzes auf die Einstellung zum Fach Chemie der HU2. Grundsätzlich wurden alle Werte für die Stichproben *real* und *angeleitet* separat ermittelt. Die Stichprobengröße bezieht sich jeweils auf die Residuenberechnung. Cronbach's Alpha wurde für den Post-Test bestimmt

Stichprobe	Gruppe	n	Pre-Test		Post-Test		Residuen		Cronbach's Alpha
			M	SD	M	SD	M	SD	
<i>real</i>	IG	81	.583	0.227	.499	0.261	.112	0.902	.94
	KG	76	.508	0.260	.409	0.257	-.120	1.082	
<i>angeleitet</i>	IG	63	.551	0.261	.462	0.275	-.049	1.003	.95
	KG	70	.534	0.254	.455	0.263	.044	0.995	

Ergänzend können anhand des Vergleichs der Post-Testergebnisse der HU1 und HU2 grundsätzliche Einflüsse der Unterrichtsreihe bzw. ihrer Dauer auf ihre Einschätzung sowohl in der IG als auch in der KG nachgewiesen werden (Tabelle 6). In allen Teilstichproben sind die Einschätzungen nach Abschluss der HU2 im Vergleich zur HU1 signifikant negativer ($IG_{real}: t(81)=3.379; p=.000; d=0.24$; $KG_{real}: t(72)=4.391; p=.000; d=0.38$; $IG_{angeleitet}: t(62)=5.803; p=.000; d=0.55$; $KG_{angeleitet}: t(69)=10.034; p=.000; d=1.05$).

8.5 Einschätzung des Aufgabeneinsatzes

Nach der HU2 unterscheiden sich nur die IG und KG der Stichprobe *real* signifikant von einander ($t(151)=-2.802; p=.006; d=0.44$), während dies in der Stichprobe *angeleitet* nicht der Fall ist ($t<1$), Tabelle 11.

Über den Vergleich der Posttestergebnisse der HU1 und HU2 lässt sich analog dem zweiten Einstellungstest ein grundsätzlicher Einfluss der Unterrichtsreihe bzw. ihrer Dauer auf die Aufgabenbeurteilung sowohl in der IG als auch in der KG nachweisen. Die Einschätzungen sind nach Abschluss der HU2 im Vergleich zur HU1 deutlich negativer (Tabelle 7 und Tabelle 11). Auffällig ist, dass nur in der Teilstichprobe IG_{real} die Einschätzung nicht signifikant schlechter ist als nach der HU1 ($IG_{real}: t(81)=1.603; p=.113; d=0.12$; $KG_{real}: t(67)=6.579; p=.000; d=0.60$; $IG_{angeleitet}: t(61)=4.776; p=.000; d=0.53$; $KG_{angeleitet}: t(66)=10.000; p=.000; d=0.90$). Dies kann als Hinweis für die motivierende Wirkung des Aufgabeneinsatzes erachtet werden, insbesondere nach der langen Beschäftigung mit dem Thema. Die Ergebnisse in der Stichprobe *real* können analog der Einschätzung zur durchgeführten Unterrichtsreihe mit einem per se positiven

Tab. 10: Wirkung des Aufgabeneinsatzes auf die Einschätzung der durchgeführten Unterrichtsreihe der HU2. Grundsätzlich wurden alle Werte für die Stichproben *real* und *angeleitet* separat ermittelt. Die Stichprobengröße und Cronbach's Alpha beziehen sich jeweils auf den Post-Test

Stichprobe	Gruppe	n	Post-Test		Cronbach's Alpha
			M	SD	
<i>real</i>	IG	83	.467	0.290	.95
	KG	74	.336	0.249	
<i>angeleitet</i>	IG	64	.355	0.287	.95
	KG	71	.338	0.275	

Tab. 11: Wirkung des Aufgabeneinsatzes auf die Einschätzung der verwendeten Aufgaben der HU2. Grundsätzlich wurden alle Werte für die Stichproben *real* und *angeleitet* separat ermittelt. Die Stichprobengröße und Cronbach's Alpha beziehen sich jeweils auf den Post-Test

Stichprobe	Gruppe	n	Post-Test		Cronbach's Alpha
			M	SD	
<i>real</i>	IG	83	.501	0.278	.93
	KG	70	.382	0.242	
<i>angeleitet</i>	IG	63	.371	0.281	.94
	KG	68	.328	0.265	

Einfluss der „Treatment-Aufgaben“ auf die Aufgaben-Einschätzung der KG, obwohl sie dieser durch den Lehrenden nur im Plenum präsentiert wurden, interpretiert werden. Insgesamt deuten die Pre/Post-Ergebnisse und sehr gute Cronbach's Alpha-Werte (aller) verwendeten Einstellungstests darauf hin, dass sich mit ihrer Hilfe die intendierten Einstellungsänderungen reliabel erfassen lassen.

9 Weitere Ergebnisse

9.1 Fähigkeitstest

Zwischen den jeweiligen Parallelklassen finden sich keine signifikanten Unterschiede. Aufgrund der Einteilung der jeweils schlechteren Parallelklasse zur IG der HU1 resultiert über alle Schulen gerechnet zwar kein signifikanter Unterschied zwischen IG und KG allerdings wird das Signifikanzniveau von .05 nur knapp überschritten ($t(239) = -1.935$; $p = .054$). Der Grund hierfür liegt – wie bereits beschrieben – in der zunächst möglichst konservativen Anlage der Untersuchung.

10 Zusammenfassung und Ausblick

Im Rahmen dieser Zusammenfassung wird sich primär auf die HU1, welche nicht durch ein vorheriges Aufgaben-Treatment beeinflusst worden ist, bezogen. In den gewählten Themenfeldern führt das systematische Bearbeiten von Aufgaben im Chemieunterricht der Sekundarstufe I zu einem signifikant höheren Lernerfolg als der herkömmliche Unterricht (Testbedingungen *real* und *angeleitet*) und entspricht damit den Vorschlägen Rosenshines, aktive Übungs- und Stillarbeitsphasen im Unterricht einzusetzen (Rosenshine, 1995). Vor allem lässt sich anhand der Testbedingung *angeleitet* der Vorteil des selbstständigen Bearbeitens gegenüber dem Bearbeiten von Aufgaben mit Hilfe der Lehrkraft explizit belegen. Da hier sowohl Interventions- als auch Kontrollgruppe mit den gleichen Aufgaben lernten,

letztere allerdings nicht eigenverantwortlich, sondern mit dem Lehrenden im Plenum, müssen die signifikanten Unterschiede zwischen beiden Gruppen primär auf die Art der Aufgabenbearbeitung zurückgeführt werden. Neben einem systematischen Aufgabeneinsatz scheint demnach insbesondere die Selbstständigkeit der Bearbeitung, die eine mentale Eigenaktivität meint, entscheidend für den Lernerfolg zu sein (Berthold & Renkl, 2008; Rosenshine, 1995).

Die selbstständige Aufgabenbearbeitung wird (nicht signifikant) positiver eingeschätzt, führt aber nicht zu einer grundsätzlichen Einstellungsänderung zum Fach Chemie. Dieses Ergebnis entspricht weiteren Studien im Fach Chemie, in denen sich eine substanzielle Einstellungsänderung weder bei relativ kurzen noch bei längeren Interventionszeiten gezeigt hat (Klos, 2009; Tepner, 2007). Hinsichtlich der Einstellung der Lernenden zur Unterrichtseinheit sind die Ergebnisse uneinheitlich. Mit zunehmender Dauer der Unterrichtsreihe wird diese von den Lernenden negativer beurteilt. Dieser Befund ähnelt den Ergebnissen von Klos, nach denen bereits im Verlauf der 7. Klasse im Fach Chemie das Sachinteresse abnimmt (Klos, 2009).

Sowohl Mädchen und Jungen als auch leistungsstärkere und leistungsschwächere Schülerinnen und Schüler profitieren in gleichem Maße vom Aufgabeneinsatz. Wechselwirkungen zwischen der Lernzeit für den Test und dem Treatment können nicht nachgewiesen werden, wie es auch von Trautwein et al. (2003) berichtet wird. Die der Konstruktion der Aufgaben und Arbeitsblätter zugrunde gelegten Gestaltungskriterien können anhand der Testergebnisse als angemessen erachtet werden und stimmen mit allgemeinen Kriterien guten Unterrichts wie z. B. klare Aufgabenstellungen und systematisches Feedback überein (Rosenshine, 1995).

Die beschriebene Studie befasst sich exemplarisch mit dem Einsatz von Aufgaben im Chemieunterricht anhand einer Unterrichtsreihe in den neunten Klassen nordrhein-

westfälischer Gymnasien. Inwieweit eine Verallgemeinerung beispielsweise auch auf andere Schulformen und -stufen, Themenbereiche oder Aufgabentypen möglich ist, müssten weitere Forschungen zeigen.

Literatur

- Ballstaedt, S.-P. (1993). Schriftliche Unterlagen für Schülerhand. *Pädagogik*, 5, 24-27.
- Becker, H.-J., Glöckner, W., Hoffmann, F., & Jüngel, G. (1992). *Fachdidaktik Chemie* (2. Aufl.). Köln: Aulis.
- Berthold, K. & Renkl, A. (2008). Wie kann eine aktive Verarbeitung von instruktionalen Erklärungen zu multiplen Repräsentationen gefördert werden? In E.-M. Lankes (Hrsg.), *Pädagogische Professionalität als Gegenstand empirischer Forschung* (S. 177-188). Münster: Waxmann.
- Bleichroth, W. (1998). Mehr Üben! *Unterricht Physik*, 9(48), 4-8.
- Bonham, S. W., Deardorff, D. L., & Beichner, R. J. (2003). Comparison of student performance using web and paper-based homework in college-level physics. *Journal of Research in Science Teaching*, 40(10), 1050-1071.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6. Aufl.). Heidelberg: Springer Medizin.
- Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, II: Effects of inequality of variance and of correlation between errors in the two-way classification. *Annals of Mathematical Statistics*, 25, 484-498.
- Brosius, F. (2004). *SPSS 12*. Bonn: MITP.
- Bruder, R. (2006). *Erläuterungen zu Modul 1 – Weiterentwicklung der Aufgabenkultur im Mathematikunterricht*. Verfügbar unter: http://sinus-transfer.uni-bayreuth.de/module/modul_1weiterentwicklung_der_aufgabenkultur.html [6.4.2007].
- Bühner, M. (2004). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung [BLK]. (1997). *Gutachten zur Vorbereitung des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“* (Materialien zur Bildungsplanung und zur Forschungsförderung Nr. 60). Bonn
- Catrambone, R. & Yuasa, M. (2006). Acquisition of procedures: The effects of example elaborations and active learning exercises. *Learning and Instruction*, 16, 139-153.
- Clauß, G., Finze, F.-R., & Partzsch, L. (2004). *Statistik für Soziologen, Pädagogen, Psychologen und Mediziner* (5. Aufl.). Frankfurt am Main: Harri Deutsch.
- Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987-2003. *Review of Educational Research*, 76(1), 1-62.
- Eckert, R. (1980). *Das Arbeitsblatt im Unterricht. Gestaltungshilfen und Beispiele für die Grund- und Hauptschule*. München: Ehrenwirth.
- Fach, M., Boer, T. de, & Parchmann, I. (2007). Results of an interview study as basis for the development of stepped supporting tools for stoichiometric problems. *Chemistry Education: Research and Practice*, 8(1), 13-31.
- Fach, M., Endres, W., & Parchmann, I. (2006). Bausteine und Stoffportionen. Erste quantitative Betrachtungen chemischer Reaktionen. *Unterricht Chemie*, 17(94/95), 26-31.
- Fach, M., Kandt, W., & Parchmann, I. (2006). Offene Lernaufgaben im Chemieunterricht. Kriterien für die Gestaltung und Einbettung. *Der mathematische und naturwissenschaftliche Unterricht*, 59(5), 284-291.
- Fechner, S. (2009). *Effects of context-oriented learning on student interest and achievement in chemistry education*. Studien zum Physik- und Chemielernen. Berlin: Logos.
- Fischer, H. E. & Draxler, D. (2001). Aufgaben und naturwissenschaftlicher Unterricht. *Der mathematische und naturwissenschaftliche Unterricht*, 54(7), 388-393.
- Hammer, C. (2002). Eigenständiges Lösen von Aufgaben. *Unterricht Physik*, 13(67), 16-17.
- Hascher, T. & Bischof, F. (2000). Integrierte und traditionelle Hausaufgaben in der Primarschule – ein Vergleich bezüglich Leistung, Belastung und Einstellungen zur Schule. *Psychologie in Erziehung und Unterricht*, 47(4), 252-265.
- Häußler, P. & Lind, G. (1998). *Erläuterungen zu Modul 1 mit Beispielen für den Physikunterricht. Weiterentwicklung der Aufgabenkultur im mathematisch-naturwissenschaftlichen Unterricht*. Verfügbar unter: http://sinus-transfer.uni-bayreuth.de/module/modul_1weiterentwicklung_der_aufgabenkultur.html [6.4.2007].
- Klos, S. (2009). *Kompetenzförderung im naturwissenschaftlichen Anfangsunterricht – Der Einfluss eines integrierten Unterrichtskonzepts*. Studien zum Physik- und Chemielernen: Bd. 89. Berlin: Logos.

- Kobayashi, K. (2006). Combined effects of note-taking/-reviewing on learning and the enhancement through interventions: A meta-analytic review. *Educational Psychology*, 26, 368-381.
- Krüger, A. (2009). Fehlersuche. *Unterricht Physik*, 14(75/76), 78-79.
- Labudde, P. (1997). Selbstständig lernen. Eine Chance für den Physikunterricht. *Naturwissenschaften im Unterricht Physik*, 8(37), 4-9.
- Mackensen-Friedrichs, I. (2005). *Förderung des Expertisenverws durch das Lernen mit Beispielaufgaben im Biologieunterricht der Klasse 9*. Dissertation, IPN. Universität Kiel. Verfügbar unter: http://eldiss.uni-kiel.de/macau/receive/dissertation_diss_00001303 [17.10.2010].
- Nentwig, P., Christiansen, D., & Steinhoff, B. (2004). Eine neue Aufgabekultur für einen neuen Chemieunterricht. *Praxis der Naturwissenschaften Chemie*, 53(8), 21-24.
- Nicolai, N. (2005). *Skriptgeleitete Eltern-Kind-Interaktion bei Chemiehausaufgaben. Eine Evaluationsstudie im Themenbereich Säure-Base*. Studien zum Physik- und Chemielernen: Bd. 43. Berlin: Logos Verlag.
- Nolen, S. B. (2003). Learning environment, motivation, and achievement in high school science. *Journal of Research in Science Teaching*, 40, 347-368.
- Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8, 293-316.
- PISA-Konsortium Deutschland (Hrsg.). (2008). *PISA 2006 in Deutschland: Die Kompetenzen der Jugendlichen im dritten Ländervergleich*. Münster: Waxmann.
- Pöpping, W. & Melle, I. (2001). Üben im Chemieunterricht. Eine Analyse von Unterrichtsbeobachtungen. *Der mathematische und naturwissenschaftliche Unterricht*, 54(7), 417-419.
- Roeder, B. (2006). Schülerbefragung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. 3. Aufl. (S. 637-642). Weinheim: Beltz.
- Rosenshine, B. (1995). Advances in research on instruction. *The Journal of Educational Research*, 88 (5), 262-268.
- Schworm, S. & Renkl, A. (2002). Lernen effektive Lösungsbeispiele zu erstellen: Ein Experiment zu einer computer-basierten Lernumgebung für Lehrende. *Unterrichtswissenschaft* (1), 7-26.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK]. (2004). *Einheitliche Prüfungsanforderungen in der Abiturprüfung Chemie (Beschluss der Kultusministerkonferenz 01.12.1989 i.d.F. vom 05.02.2004)*. Verfügbar unter: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Chemie.pdf [14.11.2010].
- Slunt, K. & Giancarlo, L. (2004). Student-centered learning: A comparison of two different methods of instruction. *Journal of Chemical Education*, 81, 985-988.
- Stäudel, L. (2003). Zäune, Geländer, Halteseile. Reflexion über Aufgaben. In H. Ball, G. Becker, R. Bruder, R. Girmes, L. Stäudel & F. Winter (Hrsg.), *Aufgaben. Lernen fördern – Selbstständigkeit entwickeln. Friedrich Jahresheft XXI 2003* (S. 18-20). Seelze: Friedrich.
- Stäudel, L. (2004a). Aufgaben für den Chemieunterricht. Eine Einführung. *Unterricht Chemie*, 15(82/83), 4-6.
- Stäudel, L. (2004b). Aufgaben nach PISA? – Aufgaben vor PISA! *Unterricht Chemie* (82/83), 91-94.
- Tepner, M. (2007). *Effektivität des Gruppenpuzzles im Chemieunterricht der Sekundarstufe I*. Dissertation, Technische Universität Dortmund. Verfügbar unter: <http://hdl.handle.net/2003/24917> [27.8.2010].
- Tepner, O. (2008). *Effektivität von Aufgaben im Chemieunterricht der Sekundarstufe I*. Studien zum Physik- und Chemielernen: Bd. 76. Berlin: Logos.
- Tepner, O. & Melle, I. (2003). Übungen im Chemieunterricht. Ergebnisse einer Befragung. *Unterricht Chemie*, 14(74), 42-43.
- Tepner, O. & Melle, I. (2004). Die Übung im Chemieunterricht – eine Analyse der aktuellen Situation. In A. Pitton (Hrsg.), *Chemie- und physikdidaktische Forschung und naturwissenschaftliche Bildung. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung der GDGP in Berlin 2003* (S. 296-298). Münster: Lit.
- Tiemann, R., Draxler, D., & Labusch, S. (2004). Aufgaben als Gegenstand empirischer Lehr-/Lernforschung. In A. Pitton (Hrsg.), *Chemie- und physikdidaktische Forschung und naturwissenschaftliche Bildung. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung der GDGP in Berlin 2003* (S. 51-53). Münster: Lit.
- Tiemann, R., Rumann, S., Jatzwauk, P., & Sandmann, A. (2006). Die Bedeutung von Aufgaben aus Lehrersicht. *Der mathematische und naturwissenschaftliche Unterricht*, 59(5), 304-308.

- Tiemann, R., Rumann, S., Jatzwauk, P., & Schabram, N. (2006). Aufgaben im Unterricht. Ansprüche – Absichten – Perzeptionen: ein Werkstattbericht. In A. Pitton (Hrsg.), *Lehren und Lernen mit neuen Medien. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung der GDChP in Paderborn 2005* (S. 266–269). Münster: Lit.
- Trautwein, U. & Köller, O. (2003). Was lange währt, wird nicht immer gut. Zur Rolle selbstregulativer Strategien bei der Hausaufgaben erledigung. *Zeitschrift für Pädagogische Psychologie*, 17(3/4), 199-209.
- Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: support for a domain-specific, multilevel homework model. *Journal of Educational Psychology*, 98(2), 438-456.
- Uihlein, A., Graf, D., & Klee, R. (2003). Vergleich zweier Aufgabentypen bei der Diagnose von Verstehensprozessen im Biologieunterricht. *Der mathematische und naturwissenschaftliche Unterricht*, 56(3), 132-136.
- von Aufschnaiter, C. & von Aufschnaiter, S. (2001). Eine neue Aufgabenkultur für den Physikunterricht. Was fachdidaktische Lernprozess-Forschung zur Entwicklung von Aufgaben beitragen kann. *Der mathematische und naturwissenschaftliche Unterricht*, 54(7), 409-416.
- Weiss, R. (1998). *Grundintelligenztest Skala 2. CFT 20* (4. Aufl.). Göttingen: Hogrefe.
- Woest, V. (2004). Aufgabenformate. *Unterricht Chemie*, 15(82/83), 7-13.

Kontakt

Dr. Oliver Tepner
 Universität Duisburg-Essen
 Didaktik der Chemie
 Schützenbahn 70
 45127 Essen
oliver.tepner@uni-due.de

Autoreninformationen

Dr. Oliver Tepner hat nach dem Studium der Fächer Chemie und Musik für das Lehramt an Gymnasien an der TU Dortmund im Arbeitskreis von Prof. Dr. Insa Melle seine Dissertation angefertigt. Sein Referendariat absolvierte er am Studienseminar Dortmund II und arbeitet nun als Akademischer Rat in der Arbeitsgruppe von Prof. Dr. Elke Sumfleth an der Fakultät für Chemie der Universität Duisburg-Essen. Seine Forschungsschwerpunkte liegen in der empirischen Lehr-Lernforschung, u. a. auf dem Professionswissen von Chemie-Lehrkräften.

Prof. Dr. Burkhard Roeder hatte bis zu seiner Emeritierung den Lehrstuhl für Methodenlehre der Psychologie an der TU Dortmund im Institut für Psychologie inne. Seine Forschungsschwerpunkte sind Lehr-Lernforschung, Geschlechtsspezifität und Einstellungsmessung.

Prof. Dr. Insa Melle ist Inhaberin des Lehrstuhls für Didaktik der Chemie II an der TU Dortmund. Ihre Forschungsschwerpunkte sind die empirische Lehr-Lernforschung und die Entwicklung von Unterrichtsmaterialien und Konzeptionen für den Chemieunterricht. Sie ist Leiterin des Chemielehrerfortbildungszentrums Dortmund.

