

Büşra Tonyalı

**Untersuchung der Bedingungen
und Wirkungen von Feedback
auf die professionelle Kompetenz
von angehenden Lehrkräften
im Vorbereitungsdienst zum
Einsatz externer Repräsentationen**

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf und Mathias Ropohl

Diese Reihe im Logos Verlag Berlin lädt Forscherinnen und Forscher ein, ihre neuen wissenschaftlichen Studien zum Physik- und Chemielernen im Kontext einer Vielzahl von bereits erschienenen Arbeiten zu quantitativen und qualitativen empirischen Untersuchungen sowie evaluativ begleiteten Konzeptionsentwicklungen zu veröffentlichen. Die in den bisherigen Studien erfassten Themen und Inhalte spiegeln das breite Spektrum der Einflussfaktoren wider, die in den Lehr- und Lernprozessen in Schule und Hochschule wirksam sind.

Die Herausgeber hoffen, mit der Förderung von Publikationen, die sich mit dem Physik- und Chemielernen befassen, einen Beitrag zur weiteren Stabilisierung der physik- und chemiedidaktischen Forschung und zur Verbesserung eines an den Ergebnissen fachdidaktischer Forschung orientierten Unterrichts in den beiden Fächern zu leisten.

Martin Hopf und Mathias Ropohl

Studien zum Physik- und Chemielernen

Band 399

Büşra Tonyalı

**Untersuchung der Bedingungen
und Wirkungen von Feedback auf die
professionelle Kompetenz von angehenden
Lehrkräften im Vorbereitungsdienst
zum Einsatz externer Repräsentationen**

Logos Verlag Berlin



Studien zum Physik- und Chemielernen

Martin Hopf und Mathias Ropohl [Hrsg.]

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Die frei zugängliche Open-Access-Publikation des vorliegenden Titels wurde mit Mitteln des Publikationsfonds der Universitätsbibliothek Duisburg-Essen ermöglicht.



Dieses Werk ist lizenziert unter der Creative Commons Lizenz CC BY-SA (<https://creativecommons.org/licenses/by-sa/4.0/>). Die Bedingungen der Creative-Commons-Lizenz gelten nur für Originalmaterial. Die Wiederverwendung von Material aus anderen Quellen (gekennzeichnet mit Quellenangabe) wie z. B. Schaubilder, Abbildungen, Fotos und Textauszüge erfordert ggf. weitere Nutzungsgenehmigungen durch den jeweiligen Rechteinhaber.

© Büşra Tonyalı

Logos Verlag Berlin GmbH 2026

ISBN 978-3-8325-6112-3

ISSN 1614-8967

DOI 10.30819/6112

Logos Verlag Berlin GmbH
Georg-Knorr-Str. 4, Geb. 10
D-12681 Berlin

Tel.: +49 (0)30 / 42 85 10 90

Fax: +49 (0)30 / 42 85 10 92

<https://www.logos-verlag.de>

Dissertation zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)
vorgelegt der Fakultät für Chemie der Universität Duisburg-Essen

**Untersuchung der Bedingungen und Wirkungen
von Feedback auf die professionelle Kompetenz
von angehenden Lehrkräften im Vorbereitungsdienst
zum Einsatz externer Repräsentationen**

von
Büşra Tonyalı
geboren in Gelsenkirchen

Mai 2025

Das von der Deutschen Forschungsgemeinschaft (DFG) geförderte Forschungsprojekt mit dem Titel „**FiRe₂ – Feedback im Referendariat zu externen Repräsentationen**“ wurde im Zeitraum von September 2018 bis September 2022 unter der Betreuung von Herrn Prof. Dr. Mathias J. Ropohl und Frau Prof. Dr. Julia Schwanewedel durchgeführt und in der vorliegenden Dissertation verschriftlicht.

1. Gutachter: Prof. Dr. Mathias J. Ropohl.

2. Gutachterin: Prof. Dr. Julia Schwanewedel

Prüfungsvorsitzender: Prof. Dr. Oliver J. Schmitz

Tag der Disputation: 17. Juli 2025

Mein Dank gilt allen, die zum Gelingen dieser Arbeit beigetragen haben: Herrn Prof. Dr. Mathias J. Ropohl und Frau Prof. Dr. Julia Schwanewedel danke ich herzlich für die Vergabe des Themas und die Betreuung meiner Arbeit. Herrn Prof. Dr. Oliver J. Schmitz danke ich für die Übernahme des Prüfungsvorsitzes. Den Verantwortlichen der Projekte KiL/KeiLa und ProwiN – Prof. Dr. Ilka Parchmann, Prof. Dr. Mirjam Steffensky und Prof. Dr. Oliver Tepner – danke ich für die Bereitstellung von Erhebungsinstrumenten zur Messung des Fachwissens und fachdidaktischen Wissens von angehenden Lehrkräften des Faches Chemie.

Dem Zentrum für Informations- und Mediendienste (ZIM) der Universität Duisburg Essen danke ich für die technische Unterstützung bei der Nutzung von Moodle sowie für die gemeinsam erstellten Abbildungen für die Intervention. Der DFG danke ich für die Finanzierung des Forschungsvorhabens, in das die Dissertation eingebunden war. Mein weiterer Dank gilt dem Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, Referat 423 (Vorbereitungsdienst, Zentren für schulpraktische Lehrerbildung, Qualifizierung im Seiteneinstieg) für die Ermöglichung der Interventionsstudie. Ebenso danke ich allen teilnehmenden Chemiefachseminaren, Fachleitungen und angehenden Lehrkräften für ihre wertvolle Zeit, Mühe und ihr Vertrauen in dieses Forschungsvorhaben.

ABSTRACT

Investigation of the conditions and effects of feedback on in-service teacher candidates' professional competence for the use of external representations

Learning chemistry involves the use of highly specialized representations, such as chemical formulas, reaction equations and illustrations of molecular models. These chemical representations are often combined and presented as multiple representations. Empirical evidence shows that working with representations is generally challenging for science learners. In particular, pre-service teachers have difficulties when working with representations and show low representational knowledge. They struggle to develop competences in evaluating and designing representations and instructional materials. Major challenges include teaching and illustrating the three levels of representations in chemistry (macro, sub-microscopic, and symbolic level). However, planning and preparing instructional materials are part of teachers' daily work. Depending on their design and use, representations can be either a facilitator or a barrier to learning. Students' learning processes are inhibited if representations are poorly illustrated or single representations are not linked. Therefore, teachers must have representational competence to prepare effective instructional materials of good quality.

Models of teachers' professional competence define subject-specific knowledge models such as content knowledge (CK) and pedagogical content knowledge (PCK). For chemistry, they explicitly include the use of representations as an essential aspect. The facets particularly relevant for the design of instructional materials are representational CK and PCK. Representational CK deals with the chemical characteristics and correctness of representations. Representational PCK addresses how chemical representations can be presented to students and how typical student misconceptions and difficulties can be minimized. In addition, teachers should have developed beliefs that consider the relevance of representations for learning chemistry. Representational beliefs can determine the extent to which teachers use their representational knowledge.

Despite this, using representations is rarely part of in-service teacher training in Germany. Regarding teacher education programs, German teachers complete in-

service teacher training after finishing their studies at university. During this time, the in-service teachers work under the supervision of mentors and experienced teachers in schools. After completing their in-service teacher training, teachers are expected to develop their professional competences and skills to the extent that they can plan and deliver lessons independently. This means in-service teacher candidates should have developed their representational competences and ability to self-improve and reflect on their competences and teaching success.

To enhance in-service teacher candidates' skills, feedback tools are an effective support. Empirical evidence confirms the high potential of feedback for teacher education and training, especially the combination of internally and externally provided feedback. Internal feedback is information that results from reflection on one's performance (e.g., a teacher candidate giving himself or herself feedback). External feedback is information someone else provides (e.g., a supervisor giving feedback to a teacher candidate). By using and combining both sources of feedback, the in-service teacher candidates can increase their representational competence, for example, for designing representations and instructional materials.

Based on the theoretical background, it is known that teacher candidates have a lack of representational knowledge, which can inhibit their ability to design instructional materials and their self-assessment skills. To minimize these deficiencies, an intervention was developed and evaluated. The research questions were: What are the effects of internal and external feedback on in-service teacher candidates' (1A) representational CK and PCK, (1B) representational beliefs, (1C) self-designed instructional materials, and (1D) representational self-assessment skills?

The intervention study was conducted in a pre-post-test design and as an interactive online learning tool. To analyze the effect of internal and external feedback, the sample ($N = 90$ 64 % = female) was randomized: The intervention group 1 (IG1, $n = 30$) used both internal and external feedback, the intervention group 2 (IG2, $n = 29$) used only internal feedback, and the control group (CG, $n = 31$) worked with dummy feedback.

Adapted test instruments measured in-service teacher candidates' representational CK and PCK before and after the intervention. A five-point Likert-scale

evaluation sheet provided internal and external feedback during the intervention. The sheet listed quality criteria for good representations and instructional materials. After designing instructional material, the in-service teacher candidates used the sheet to assess their instructional material (internal feedback). Subsequently, the researchers also assessed the in-service teacher candidates' instructional materials, and the in-service teacher candidates received that rating as external feedback for further reflection. While the intervention groups worked with systematic and differentiated feedback, the dummy feedback group received only a simple, superficial, non-individual, short text as external feedback. The interrater reliability between the internal and external feedback was used to quantitatively determine the in-service teacher candidates' self-assessment skills (high reliability = correct rating/assessment).

The whole intervention was divided into three structurally identical sessions. In each session, the in-service teacher candidates were given an exercise to design a worksheet and a specific learning objective for their students. First, they generated internal feedback; second, they received external feedback; and lastly, they compared both feedback information to reflect and identify possibilities to improve the instructional materials. This cycle of exercise, feedback, and reflection was repeated three times.

The results revealed differential effects in the representational PCK and self-assessment skills. The residuals show a significant improvement in PCK of the IG2 ($\eta^2 = .07$), whereas the growth in CK did not differ between the groups. There were also no substantial changes in their representational beliefs or designed instructional materials in time or group comparison.

Looking at the effect of feedback on representational self-assessment skills in the IG1 and IG2, both groups improved their skills. However, there was a significant group effect ($\eta^2 = .35$): In the first session, the self-assessment skills of both groups were almost equal. In the second session, a group effect was only descriptive. The IG1 was minimally better than the IG2. This variance significantly increased in the third session with a large effect size ($\eta^2 = .22$). Group-specific post-hoc analyses between the measurement points showed a significant improvement in the IG1 from one session to the next (session 1/2: $\eta^2 = .11$; session 2/3: $\eta^2 = .18$; session 1/3: $\eta^2 = .31$). In the IG2, the improvements were only descriptive. The improvement was only significant between the first and last

session ($\eta^2 = .07$). In addition, data on the cognitive load during the intervention showed that the IG1 had a higher cognitive load.

The results indicate differential effects of feedback: Working with criteria-based internal feedback increases the in-service teachers' representational PCK. Working with both internal and external feedback increases their representational self-assessment skills. Providing internal and external feedback more frequently promotes the skills of self-assessment. The combination has no benefit for knowledge growth, probably due to in-service teacher candidates' cognitive overload. Any source of used feedback (internal or external) cannot support CK. One possible reason why there were no changes in the beliefs or the instructional material quality could be that they were already at a high level in the pre-test.

The results are consistent with prior empirical studies about in-service teacher candidates' professional competence. They indicate the growth of practical and application-related skills (planning, implementing, reflecting on teaching) rather than the growth of knowledge (here: CK, PCK) during the in-service teacher training.

Since teacher candidates should learn to plan and conduct lessons autonomously and should also be able to improve themselves independently, these findings provide a helpful tool to enable in-service teacher candidates to work independently on giving themselves effective and precise feedback. Incorporating internal and/or external feedback into in-service teacher trainings and applying it to other relevant topics, such as designing instructional materials (e.g., planning of lessons), would become a powerful supportive tool for in-service teacher candidates. However, it should be taken into account that additional external feedback can lead to cognitive overload. Nevertheless, if both sources of feedback are used in combination, it is important to ensure that the feedback is appropriately designed with regard to in-service teacher candidates' needs for support, cognitive demands and proficiency levels (e.g., extent, didactic content and chemical content).

INHALTSVERZEICHNIS

ABSTRACT	VII
1 EINLEITUNG.....	1
2 THEORETISCHER HINTERGRUND.....	5
2.1 Multiple externe Repräsentationen in der Chemie	5
2.1.1 Relevanz und kognitive Verarbeitung von Repräsentationen	5
2.1.2 Repräsentationsformen und -ebenen in der Chemie.....	9
2.1.3 Lernen chemischer Konzepte mithilfe von Repräsentationen und einhergehende Schwierigkeiten	12
2.1.4 Kognitionspsychologische und fachdidaktische Gestaltungsprinzipien	15
2.2 Repräsentationsbezogene Kompetenz von Chemielehrkräften	20
2.2.1 Strukturierung der repräsentationsbezogenen Kompetenz	20
2.2.2 Einordnung in die professionelle Handlungskompetenz	22
2.2.3 Entwicklung der Kompetenz im Rahmen der Lehrkräftebildung	25
2.3 Feedback zur Unterrichtsplanung von Chemielehrkräften.....	28
2.3.1 Informatives tutorielles Feedback.....	29
2.3.2 Implikationen zur Feedbackgestaltung für (angehende) Lehrkräfte.....	33
2.4 Zusammenfassung des bisherigen Forschungsstandes und Desiderate.....	40
3 ZIEL, FORSCHUNGSFRAGEN UND HYPOTHESEN	43
4 METHODISCHES VORGEHEN	49
4.1 Studiendesign.....	49
4.1.1 Untersuchungsvariablen.....	49
4.1.2 Akquise und Gruppenzuordnung der Stichprobe	52
4.1.3 Struktur und Inhalt der drei Online-Interventionssitzungen.....	53
4.1.4 Entwicklung und Fachlicher Inhalt des Rohmaterials für das Lehr-Lern-Material	58

4.2	Testinstrumente und Pilotierungen	62
4.2.1	Demografische Daten und persönliche Ressourcen	62
4.2.2	Allgemeines Fachwissen und allgemeines fachdidaktisches Wissen.....	64
4.2.3	Repräsentationsbezogenes Fachwissen und fachdidaktisches Wissen.....	67
4.2.4	Repräsentationsbezogene Überzeugungen	74
4.2.5	Repräsentationsbezogene Qualität der Lehr-Lern-Materialien	75
4.2.6	Selbsteinschätzungsfähigkeit/Qualität des internen Feedbacks	81
4.3	Durchgeführte Zusammenhangsanalysen.....	83
4.3.1	Dokumentation der deskriptiven Statistik und Testkennwerte.....	83
4.3.2	Varianzanalysen.....	84
4.3.3	Bivariate Zusammenhangsanalysen (Korrelationen).....	87
4.3.4	Multivariate Zusammenhangsanalysen (Multiple lineare Regressionen).....	88
5	ERGEBNISSE DER HAUPTSTUDIE	91
5.1	Stichprobenmerkmale der angehenden Lehrkräfte.....	91
5.1.1	Demografische Daten	91
5.1.2	Persönliche Ressourcen im Vorbereitungsdienst	94
5.1.3	Erfahrungen in der Gestaltung von Lehr-Lern-Materialien	97
5.1.4	Allgemeines Fachwissen und fachdidaktisches Wissen.....	100
5.2	Repräsentationsbezogenes Fachwissen und fachdidaktisches Wissen....	103
5.3	Repräsentationsbezogene Überzeugungen	109
5.4	Qualität der designten Lehr-Lern-Materialien.....	115
5.5	Selbstbewertungsfähigkeit (Qualität des internen Feedbacks).....	123
5.5.1	Analyse der Eignung der fünfstufigen Skalierung zur Selbsteinschätzung.....	127
5.5.2	Wahrgenommener Cognitive Load während der Feedbackphasen	129
5.5.3	Wahrgenommene Anstrengungsbereitschaft während der Feedbackphasen	132
5.6	Zusammenhangsanalysen zu den Stichprobenmerkmalen	134
5.6.1	Varianzen in den demographischen Daten der Gesamtgruppe.....	134

5.6.2	Korrelationen in der Gesamtgruppe zwischen allen Variablen	138
5.6.3	Gruppenspezifische Median-Splits im Fachwissen und fachdidaktischen Wissen.....	146
5.6.4	Prädiktoren des Fachwissens und fachdidaktischen Wissens	151
6	DISKUSSION UND INTERPRETATION DER ERGEBNISSE	155
6.1	Methodische Reflexion	155
6.1.1	Güte der Testinstrumente	155
6.1.2	Eignung des Studiendesigns und Umsetzung der Intervention	156
6.2	Diskussion und Beantwortung der Forschungsfragen	159
6.2.1	F _{1A} : Effekt von internem und/oder externem Feedback auf das repräsentationsbezogene Fachwissen und fachdidaktische Wissen von angehenden Lehrkräften.....	159
6.2.2	F _{1B} : Effekt von internem und/oder externem Feedback auf die Überzeugungen der angehenden Lehrkräfte in Bezug auf die Relevanz von externen Repräsentationen für das Lernen von Chemie	160
6.2.3	F _{1C} : Effekt von internem und/oder externem Feedback auf die Qualität der von den angehenden Lehrkräften designten multiplen externen Repräsentationen	161
6.2.4	F _{1D} : Effekt von internem und/oder externem Feedback auf die Selbstbewertungsfähigkeit von angehenden Lehrkräften in Bezug auf selbstdesignte multiple externe Repräsentationen	162
6.2.5	F ₂ : Effekt der Stichprobenmerkmale der angehenden Lehrkräfte auf das repräsentationsbezogene Fachwissen und fachdidaktische Wissen.....	163
6.3	Gesamtfazit, Implikationen für die Forschung und Schulpraxis	165
6.4	Limitationen und Ausblick	167
7	VERZEICHNISSE	169
7.1	Literaturverzeichnis	169
7.2	Abkürzungsverzeichnis.....	190
7.3	Tabellenverzeichnis	192
7.4	Abbildungsverzeichnis	197

8	ANHANG.....	201
----------	--------------------	------------

1 EINLEITUNG

Das Bildungssystem steht derzeit vor vielfältigen Herausforderungen, die auch die Professionalisierung von Lehrkräften betreffen. Neben den Auswirkungen der Coronavirus-Pandemie und von Fluchtbewegungen, die zu größerer sprachlicher Heterogenität in deutschen Schulklassen führen, wirken sich insbesondere aktuelle gesellschaftliche und strukturelle Krisen auf das Bildungssystem aus. Hierzu zählen beispielsweise der anhaltende Fachkräftemangel sowie die wachsenden Anforderungen an die Digitalisierung und Inklusion im Schulwesen (Walm & Wittek, 2023).

In diesem Kontext nehmen die MINT-Fächer eine besondere Rolle ein, da der Bedarf an Lehrkräften in diesen Bereichen besonders hoch ist (Güldener, 2023). Diese Entwicklungen stellen angehende Lehrkräfte, insbesondere in den naturwissenschaftlichen Unterrichtsfächern, vor große Herausforderungen. Sie müssen nicht nur fachlich kompetent sein, sondern auch über umfassende fachdidaktische und pädagogische Fähigkeiten verfügen, um den komplexen Anforderungen eines sich wandelnden Bildungssystems gerecht zu werden.

In den naturwissenschaftlichen Unterrichtsfächern und insbesondere in der Chemie ist das Wissen über Repräsentationen auf Schüler:innenebene für den Aufbau konzeptionellen Wissens von grundlegender Bedeutung. Durch externe Repräsentationen können nicht nur chemische Phänomene abgebildet, sondern auch erklärt werden (Treagust, 2018). Studien zeigen jedoch, dass Lernende häufig Schwierigkeiten beim Verständnis und der Verwendung externer Repräsentationen haben (Cheng & Gilbert, 2009). Umso wichtiger ist es daher, dass angehende Lehrkräfte auch auf diese fachlichen Herausforderungen vorbereitet und durch konkrete Maßnahmen unterstützt werden. Befunde deuten sogar darauf hin, dass angehende Lehrkräfte selbst Schwierigkeiten mit Repräsentationen aufweisen (Taskin et al., 2015a, 2015b; Taskin & Bernholt, 2012). Dies deutet auf einen Bedarf einer repräsentationsbezogenen Unterstützungsmaßnahme für angehende Lehrkräfte hin.

Im Rahmen der Lehrkräfteprofessionalisierung wird auch erforscht, wie Lehrkräfte aus- und fortgebildet werden können. Beispielsweise existieren Erkenntnisse zur Wirksamkeit von Fortbildungsmaßnahmen für Lehrkräfte (z. B.

Lipowsky, 2011). Verschiedene Studien betonen, dass insbesondere die Verbindung von fachbezogenem Professionswissen, also Fachwissen und fachdidaktischem Wissen, sowie die professionellen Überzeugungen entscheidende Einflussfaktoren auf die Unterrichtsqualität darstellen (Kunter et al., 2011; Oser & Blömeke, 2012). Es ist jedoch unklar, inwieweit solche Befunde auf angehende Lehrkräfte im Vorbereitungsdienst übertragbar sind.

Bei der gezielten Förderung der professionellen und repräsentationsbezogenen Kompetenz kann der Einsatz von Feedback ein zentraler Hebel sein: Feedback wird als äußerst wirkungsvolles Instrument zur Unterstützung von Lernprozessen betrachtet, wobei seine Effektivität in hohem Maße von seiner Art und Quelle abhängt (Wisniewski, Zierer & Hattie, 2019):

*„Feedback, on average, is powerful, but some feedback is more powerful.“
(Wisniewski et al., 2019, S. 13)*

Während in der Regel jedoch das Feedback *durch* Lehrkräfte untersucht wird (Hattie & Timperley, 2007), wird der Einfluss von kollegialem oder tutoriellem Feedback in der Lehrkräftebildung bislang noch zu wenig berücksichtigt. Dabei kann Feedback nicht nur Auswirkungen auf das Professionswissen von Lehrkräften haben, sondern auch ihre Überzeugungen verändern – und somit letztlich auch ihr professionelles Handeln im Unterricht. Vor diesem Hintergrund stellt sich die Frage, inwieweit angehende Lehrkräfte von internem Feedback (Selbstfeedback) und/oder externem Feedback (Fremdfeedback) profitieren und wie diese Rückmeldungen in den Lernprozess von angehenden Lehrkräften integriert werden sollten, um den größten Lernerfolg zu erzielen. Es wird angenommen, dass die Wahrnehmung der Effektivität von Feedback eine entscheidende Rolle spielt und dass eine Kombination von internem und externem Feedback besonders förderlich für den Lernprozess ist (Brinko, 1993; Butler & Winne, 1995). Allerdings sind die Bedingungen und Wirkungen von Feedback im Vorbereitungsdienst noch nicht ausreichend untersucht und die bestehenden Studien zeigen teils inkonsistente Ergebnisse (Kapitel 2.3).

Um für die geschilderte Herausforderung angehender Chemielehrkräfte repräsentationsbezogene und feedbackgestützte Unterstützungsmaßnahmen zu entwickeln und erproben, wurde die vorliegende Forschungsarbeit durchgeführt.

Die Studie zielt darauf ab, das Professionswissen und die Überzeugungen von Lehramtsanwärter:innen im Umgang mit externen Repräsentationen zu fördern und zu verändern. Die gewonnenen Erkenntnisse dieser Arbeit können genutzt werden, um langfristig die zweite Phase der Lehrkräftebildung kontinuierlich zu verbessern.

Um das Forschungsanliegen systematisch zu untersuchen, wurde eine empirische und quantitative Studie durchgeführt, welche in der vorliegenden Dissertation dokumentiert wird. Die Dissertation ist klassisch und nach den Konventionen einer empirischen Forschungsarbeit der Chemiedidaktik aufgebaut. Zunächst werden in Kapitel 2 die Eigenschaften sowie der aktuelle Forschungsstand zu multiplen externen Repräsentationen, der Ausbildung von Chemielehrkräften im Vorbereitungsdienst und geeigneten Feedbackansätzen dargestellt. Darauf aufbauend werden in Kapitel 3 die konkreten Forschungsfragen und Hypothesen formuliert. Im anschließenden Kapitel 4 wird das Studiendesign der feedbackgestützten Intervention beschrieben. In Kapitel 5 werden die Ergebnisse der Hauptstudie berichtet und analysiert. Abgeschlossen wird die Arbeit in Kapitel 6 mit einer Zusammenfassung der zentralen Befunde und den daraus abgeleiteten Implikationen für die Praxis und zukünftige Forschung.

2 THEORETISCHER HINTERGRUND

2.1 Multiple externe Repräsentationen in der Chemie

2.1.1 Relevanz und kognitive Verarbeitung von Repräsentationen

Repräsentationen sind Wiedergaben von originalen Objekten, Phänomenen oder Prozessen mit unterschiedlichem Abstraktionsgrad (Ainsworth, 2006; Schnotz, 2014). Diese können beispielsweise ein originalgetreues Foto eines Versuchsaufbaus oder ein abstrahiertes Molekülmodell eines d-Orbitals sein.

Kognitionspsychologisch betrachtet, spielen Repräsentationen beim Lernen eine wichtige Rolle. Lerntheorien bauen auf dem Ansatz auf, dass neue Informationen (externe Repräsentationen) von Lernenden sensorisch aufgenommen und kognitiv verarbeitet werden, sodass in ihrer Kognition neues Wissen konstruiert wird (interne Repräsentationen). Zu den Theorien zählen beispielsweise die kognitive Theorie des multimedialen Lernens (Mayer, 2014a) sowie das integrative Modell des Text- und Bildverstehens (Schnotz, 2014), welche auf essenzielle Lerntheorien des Kognitivismus und Konstruktivismus zurückgreifen (z. B. Baddeley et al., 2015; Clark & Paivio, 1991; Piaget, 1948; Sweller et al., 2011; Wittrock, 1989).

Relevant ist die Unterscheidung zwischen abbildenden bzw. depiktionalen (bildlichen) und beschreibenden bzw. deskriptiven (textlichen) Repräsentationen. Deskriptive Repräsentationen bestehen aus Symbolen, die durch eine willkürliche Struktur gekennzeichnet sind. Konventionen legen fest, welches Objekt ein Symbol repräsentiert. Typische Beispiele sind Texte und mathematische Gleichungen. Im Gegensatz dazu bestehen bildliche Repräsentationen aus ikonischen Zeichen. Während realistische Bilder (z. B. Fotografien, Zeichnungen) dem dargestellten Objekt ähnlich sind, sind logische Bilder (z. B. Graphen, Diagramme) eher abstrakt (Schnotz, 2002a).

Anhand der Annahmen der Lerntheorien kann nicht nur abgeleitet werden, wie externe Repräsentationen bzw. Informationen verarbeitet werden, sondern auch durch welche Eigenschaften und Gestaltungen der Repräsentationen der Lernprozess unterstützt werden kann. Grundlegend kann festgehalten werden, dass Lernen aus Text-Bild-Kombinationen effektiver ist als das Lernen mit nur einem Text oder einer Abbildung (Eilam & Poyas, 2008; Levie & Lentz, 1982; Mayer, 2009a). Eine Erklärung bietet beispielsweise das integrative Modell des Text-

und Bildverstehens, welches das Lernen auf kognitiver Ebene veranschaulicht (Abbildung 1; Schnotz, 2014): Wenn Lernenden Repräsentationen dargeboten werden, erfolgt zunächst eine sensorische Aufnahme der Informationen und anschließend eine Selektion und Überführung in das Arbeitsgedächtnis. Dabei werden alle Gedächtnisleistungen in zwei Subsysteme unterteilt, da Texte und Bilder aufgrund ihrer deskriptiven und depiktionalen Struktur unterschiedlich aufgenommen und verarbeitet werden. Dabei wird der Text erst im deskriptiven Subsystem verarbeitet und erst anschließend mit dem depiktionalen Subsystem abgeglichen. Für Bilder erfolgt der Prozess in umgekehrter Weise. Beim abschließenden Aufbau des mentalen Modells bzw. der internen Repräsentation wird das neu gelernte Wissen in das Langzeitgedächtnis integriert. Letztlich werden beide externen Repräsentationen getrennt verarbeitet, gehen aber ineinander über und ergänzen sich gegenseitig. Ist im Lernprozess nur eine Variante von Repräsentationen enthalten (d. h. nur Text oder Bild), so kann kein gegenseitiger Abgleich und keine Ergänzung stattfinden. Das entstandene Modell wäre somit nur bedingt ausgereift und vollständig (Schnotz, 2014).

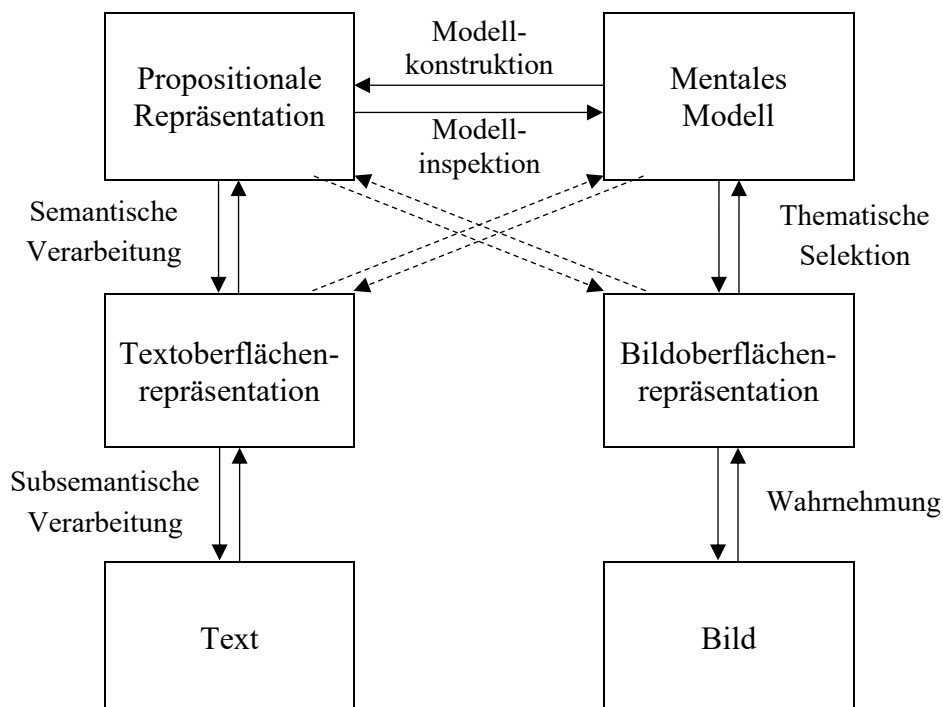


Abbildung 1: Modell des integrativen Text- und Bildverstehens (Schnotz, 2014; Schnotz et al., 2022)

In der Chemie ist es üblich, verschiedene Formen externer Repräsentationen zu kombinieren. Chemische Phänomene oder Prozesse sind meist so komplex und vielfältig, dass sie durch eine Repräsentationsform nicht ausreichend abgebildet werden können (Permatasari et al., 2022). Unterkategorien, wie Informationstexte, Versuchsanleitungen, chemische und mathematische Formeln, aber auch Teilchenabbildungen und verschiedene Arten von Diagrammen helfen, diese chemischen Phänomene und Prozesse aus möglichst vielen Perspektiven zu beleuchten (Tippett, 2016; Treagust et al., 2017). Beispielsweise ist die verbale Erklärung von Aggregatzuständen ohne Einsatz von Molekülabbildungen schwer vorstellbar, genauso wie das Darstellen einer Säure-Base-Titration ohne zugehörige Reaktionsgleichungen und Titrationskurven. Durch derartige Kombinationen von mindestens zwei Repräsentationsformen zum selben oder ähnlichen Sachverhalt entstehen multiple externe Repräsentationen (Ainsworth, 2006, 2018; Seufert, 2003). Im Folgenden werden diese Kombinationen und deren Eigenschaften anhand von Beispielen aus der Chemie veranschaulicht.

Eine funktionale Klassifikation von multiplen externen Repräsentationen, welche auf den benannten Lerntheorien aufbaut, erfolgt in den Arbeiten von Ainsworth (Abbildung 2; z. B. Ainsworth, 1999, 2006, 2018). In Abhängigkeit zu ihrer konkreten Gestaltung können multiple Repräsentationen die Lernenden abholen, indem sie an ihr Vorwissen anknüpfen und sie nicht durch ihre Komplexität überfordern. Lehrkräfte stehen somit vor der Aufgabe, multiple Repräsentationen auf einem angemessenen Anspruchsniveau zu präsentieren, so dass einerseits die kognitive Belastung möglichst gering ist und andererseits Lernprozesse auf einem förderlichen Niveau aktiviert werden. Folgende Merkmale können zur Änderung des Anforderungsmaßes variiert werden (Ainsworth, 2006):

- die Formen der eingesetzten Repräsentationen (z. B. Text, Formel, Diagramm, Abbildung),
- die Anzahl der kombinierten Repräsentationen,
- die Schnittmenge an Informationen, die die einzelnen Repräsentationen liefern,
- die Reihenfolge bzw. Anordnung der Repräsentationen,
- die Funktion der Verknüpfung der einzelnen Repräsentationen untereinander.

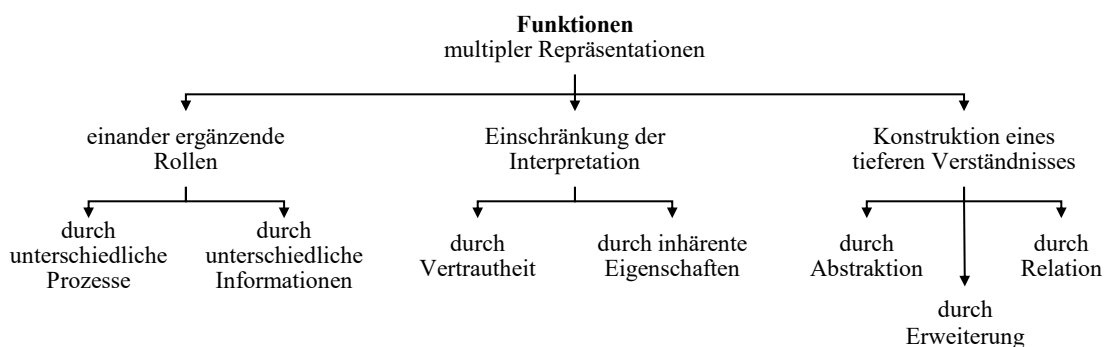


Abbildung 2: Funktionale Taxonomie multipler externer Repräsentationen (Ainsworth, 1999; übersetzt nach Krey & Schwanewedel, 2018)

Ein Parameter, der eine wesentliche Rolle bei der Zusammenführung mehrerer einzelner Repräsentationen zu einer multiplen Repräsentation spielt, ist die Funktion der Verknüpfung. In der funktionalen Taxonomie multipler Repräsentationen werden drei Arten unterschieden, wie einzelne Repräsentationen miteinander inhaltlich wechselwirken können (Ainsworth, 1999). Maßgeblich für die Kategorisierung ist der Grad der inhaltlichen Redundanz: Ergänzung (hoch), Einschränkung (partiell), Konstruktion (gering). Einzelne Repräsentationen können komplementäre Informationen enthalten, um beispielsweise die Komplexität eines anspruchsvollen Lerngegenstands zu reduzieren. Je nach Jahrgangsstufe und Lernziel kann dies zum Beispiel die Einführung des einfachen Teilchenmodells sein, indem eine Abbildung das Modell zeigt und im Text die Annahmen des Modells erläutert werden. Bei der Einschränkungsfunktion dient eine Repräsentation dazu, die Interpretation einer zweiten Repräsentation zu lenken (Ainsworth, 2006). Dies ist beispielsweise der Fall, wenn eine Repräsentationsform den Lernenden bereits vertraut ist (z. B. Strukturformeln) und dabei hilft, eine neue Repräsentationsform (z. B. die Nomenklatur der organischen Chemie) einzuführen. Im dritten Fall können multiple Repräsentationen bei einem bereits bestehenden Grundverständnis der Lernenden eingesetzt werden, um das Wissen mit einem tieferen Verständnis zu ergänzen (Ainsworth, 2006). Eine Erweiterung liegt beispielsweise vor, wenn das chemische Gleichgewicht eingeführt werden soll. Ein klassischer Versuch hierzu erfolgt, indem Lernende eine Gleichgewichtsreaktion durchführen, danach die Konzentrationsveränderungen in Abhängigkeit zum Zeitverlauf in einer Tabelle auflisten und die Daten anschließend auf einem Punkt-Liniendiagramm analysieren.

2.1.2 Repräsentationsformen und -ebenen in der Chemie

In Bezug auf den Einsatz und Gebrauch von Repräsentationen, weist das Fach Chemie, wie auch die anderen naturwissenschaftlichen Fächer, eine Besonderheit auf (Permatasari et al., 2022). In der Chemie finden einerseits Repräsentationen Anwendung, die einen hohen Alltagsbezug aufweisen oder fachunabhängig genutzt werden (z. B. Texte, Fotografien) und mehrheitlich in weiteren naturwissenschaftlichen Disziplinen zum Einsatz kommen (z. B. Diagramme, mathematische Formeln). Andererseits wird die Chemie als ein abstraktes und schwieriges Fachgebiet wahrgenommen, in dem eine Vielzahl von facheigenen, komplexen Repräsentationsformen zum Einsatz kommt (Gilbert & Treagust, 2009b; Hoffmann & Laszlo, 1991; Santos & Arroio, 2016). Beispielsweise werden Diagramme nicht nur verwendet, um den Zusammenhang zwischen zwei Variablen zu veranschaulichen (z. B. Phasendiagramm, Energiediagramm), sondern auch, um energetische Verhältnisse in organischen Molekülen darzustellen (z. B. Frost-Muslin-Diagramm). Daneben stellen chemische Formeln eine Repräsentationsform dar, die so fachtypisch, detailorientiert und markant ist, dass sie aufgrund ihrer Komplexität und Besonderheiten als eine Art „Sprache für sich selbst“ betrachtet werden können (Fluck, 1996; Hoffmann & Laszlo, 1991). Generell lässt sich festhalten, dass Repräsentationen mit dem steigenden fachlichen Anspruch der chemischen Inhalte spezifischer und komplexer werden (Hoffmann & Laszlo, 1991). Eine zusammenfassende Übersicht über die in der Chemie verwendeten Repräsentationsformen ist in Abbildung 3 zu finden.

Je nachdem welche Repräsentationsform ausgewählt wird, um ein Original (Objekt, Phänomen oder Prozess) abzubilden, ändert sich der Abstraktionsgrad. So können die Repräsentationen, je nach ihrer Funktion, das Original unterschiedlich stark abstrahieren oder nur bestimmte Aspekte des Originals wiedergeben (Treagust & Tsui, 2013a). Das Foto einer Blauen Lotusblume in Abbildung 4 hat zum Beispiel die größte Ähnlichkeit mit der Originalblüte. Form, Farben und Größenverhältnisse der Blüte und Blütenbestandteile sowie einige Blätter und ein Teil des Sees sind zu erkennen. Das (Blüten-)Diagramm enthält hingegen nur die spezifischen Merkmale der Lotusart und kann daher jeden Blaue Lotus im Allgemeinen darstellen. Eine Gleichung bzw. mathematische Formel, die die

Höhe der Blume in Abhängigkeit zum Zeitverlauf angibt, steht in keiner ikonischen Beziehung zum Original. Dennoch drückt die Formel ein Merkmal der Pflanze bzw. ihre Wachstumshöhe aus, auch wenn der Abstraktionsgrad so hoch ist, dass der Zusammenhang nicht direkt erkennbar ist (Roth & Pozzer-Ardenghi, 2013).

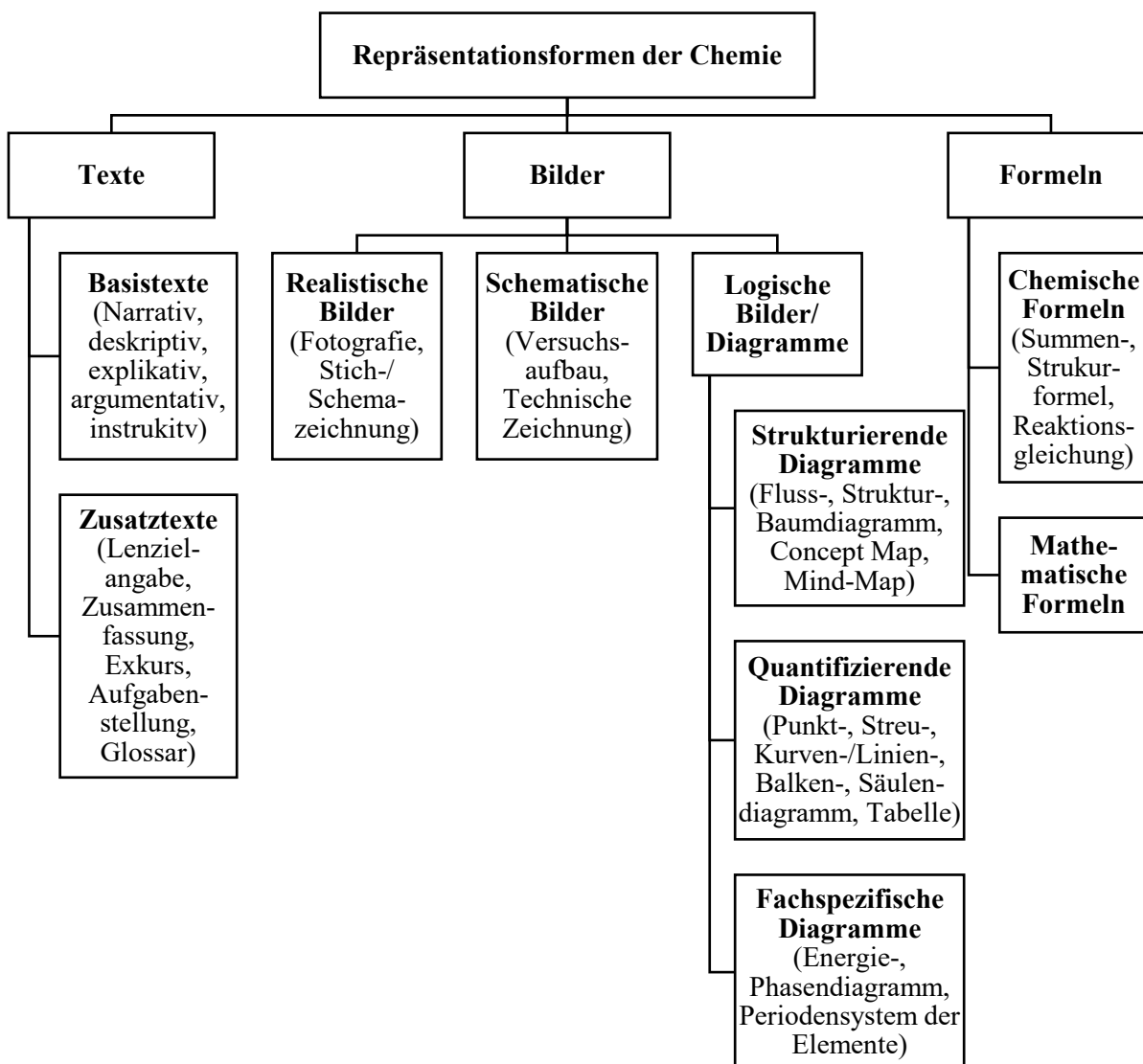


Abbildung 3: Repräsentationsformen in der Chemie mit ausgewählten Beispielen (erstellt nach Gilbert & Treagust, 2009b; Lachmayer et al., 2007; Leisen, 2005; Mangold, 2015; Nerdel, 2017; Peterson, 1996; Treagust et al., 2017; Treagust & Tsui, 2013b; Weidenmann, 1988)

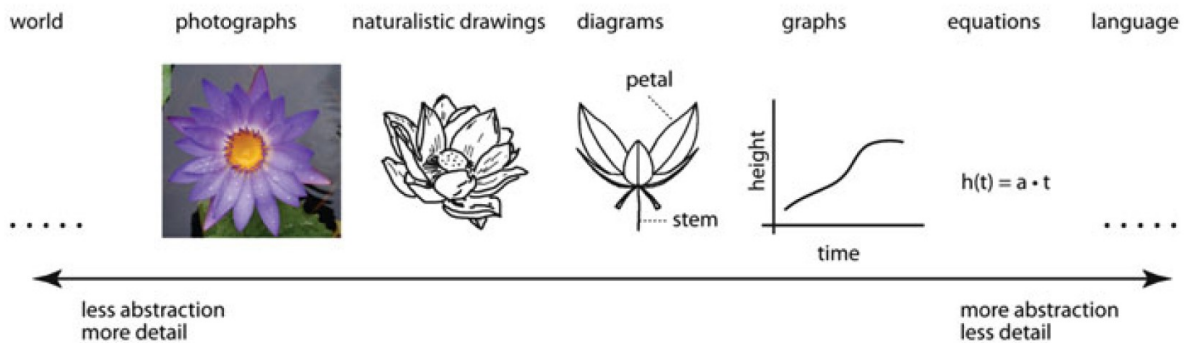


Abbildung 4: Abstraktionsgrade von externen Repräsentationen in den Naturwissenschaften am Beispiel des Blauen Lotus (verändert nach Roth & Pozzer-Ardenghi, 2013)

Ein weiteres fachspezifisches Merkmal multipler Repräsentationen sind die drei Repräsentationsebenen der Chemie (Permatasari et al., 2022). Lehrkräfte nutzen die Ebenen im Unterricht, um den Schüler:innen die chemischen Phänomene aus unterschiedlichen Perspektiven nahezubringen. Welche Ebene sich dabei eignet, hängt vom Leistungstand der Lernenden und dem konkreten Lernziel ab. Dabei wird zwischen der makroskopischen bzw. phänomenologischen Ebene, der submikroskopischen bzw. Modell-Ebene sowie der symbolischen Ebene unterschieden (Gilbert & Treagust, 2009a). Bei der Entwicklung eines Verständnisses von naturwissenschaftlichen Konzepten ist es aus fachlicher Sicht notwendig, zwischen diesen verschiedenen Abstraktionsebenen zu unterscheiden. Ebenen von Repräsentationen geben an, aus welcher Perspektive das Phänomen oder der Prozess dargestellt wird. Diese Klassifizierung geht auf Johnstone (2000) zurück und wurde mehrfach neu interpretiert und angepasst (z. B. Farida et al., 2018; Talanquer, 2011).

Auf der *makroskopischen* Ebene werden Objekte und Prozesse dargestellt, die mithilfe der Sinneswahrnehmung registriert werden können. Dazu zählen Stoffe und Objekte, wie Feststoffe, Flüssigkeiten, Gase, Kolloide, Aerosole und deren Eigenschaftsveränderungen, wie Farbumschläge und Lösungsvorgänge. Ebenso werden auf der makroskopischen Ebene die messbaren Eigenschaften von Stoffen dargestellt, wie die Masse, Dichte, Konzentration und Temperatur (Gilbert & Treagust, 2009a). Prozesse und Phänomene, bei denen die Informationen auf der makroskopischen Ebene nicht erklärbar sind, können durch die *submikroskopische* Ebene modellhaft dargestellt und dadurch erklärt werden. Oft werden

dabei Visualisierungen von Atomen, Molekülen, Ionen und Partikeln genutzt (Davidowitz & Chittleborough, 2009; Gilbert & Treagust, 2009a). Die *symbolische* Ebene kann als Mediator zwischen der makroskopischen und submikroskopischen Ebene fungieren (Taber, 2009). Chemische Darstellungen, die der makroskopischen (z. B. Zugabe von Kochsalz in Wasser) oder submikroskopischen Ebene (hier: Kugelteilchenmodell von Wasser und Natriumchlorid) angehören, können auch auf der symbolischen Ebene (hier: Reaktionsgleichung des Lösevorgangs) dargestellt werden. Beschreibungen und Abkürzungen der symbolischen Ebene sind durch Konventionen der Chemie und Mathematik genau definiert. Dies trifft beispielsweise Reaktionsgleichungen und darin enthaltene Elementsymbole und stöchiometrische Koeffizienten, Ladungszahlen, Formelzeichen (Gilbert & Treagust, 2009a; Taber, 2009). Eine multiple externe Repräsentation in diesem Sinne könnte z. B. die Kombination zweier Bilder in zwei verschiedenen Ebenen sein: ein Foto eines Eisenstücks (makroskopische Ebene) und ein Bild der Gitterstruktur der Eisenatome (submikroskopische Ebene).

2.1.3 Lernen chemischer Konzepte mithilfe von Repräsentationen und einhergehende Schwierigkeiten

Auf Grundlage ihrer Untersuchungen zum Lernen mit multiplen externen Repräsentationen haben Kozma und Russel (1997) für den Chemieunterricht Schlussfolgerungen für Lehrpläne und Unterricht gezogen. So sollten die Schüler:innen beispielsweise in der Lage sein:

- Elemente, Strukturen und Muster chemischer Phänomene in einer bestimmten Repräsentation zu erkennen und zu analysieren,
- eine Repräsentationsform in eine andere umzuwandeln,
- eine oder mehrere geeignete Repräsentationen auszuwählen oder zu erzeugen um chemische Phänomene oder Konzepte zu erklären.

Die Schwierigkeiten von Schüler:innen beim Arbeiten mit Repräsentationen basieren jedoch nicht nur auf fachlichen Wissens- und Kompetenzlücken, sondern beginnen bereits mit der mangelnden Auseinandersetzung und Enkodierung der dargebotenen Repräsentationen und Lernmaterialien. Befunde deuten darauf hin, dass Lernende sich nicht detailliert genug mit Lernmaterial auseinandersetzen, wodurch nur ein eingeschränkter Lernzuwachs möglich ist. Eine dieser

Schwierigkeiten besteht darin, dass Schüler:innen oft nur Elemente oder Strukturen auf der Oberflächenebene einer Repräsentation wahrnehmen, wie Symbole, Farben oder Linien. Weiterführende Informationen in den Repräsentationen, die beispielsweise genutzt werden können, um Annahmen zu untermauern oder zu erklären, Schlüsse zu ziehen und Vorhersagen über Beziehungen zwischen chemischen Phänomenen oder Konzepten zu treffen, werden von den Lernenden nicht verarbeitet (Bodner & Domin, 2000).

Darüber hinaus zeigen Cromley et al. (2010), dass Schüler:innen das Potenzial naturwissenschaftlicher Bilder nicht vollständig nutzen. Lernende, die mit Texten und Bildern zugleich arbeiten, betrachten die Bilder flüchtig oder missachten diese gänzlich. Lernende, die sich jedoch ausreichend auf die Bilder konzentrieren, können sich mit diesen auf einem höheren Niveau und mit einer höheren kognitiven Aktivität auseinandersetzen. Es ist daher davon auszugehen, dass Lernende erst ein Verständnis für naturwissenschaftliche Konzepte entwickeln können, wenn sie aktiv an und mit multiplen externen Repräsentationen arbeiten und dabei naturwissenschaftliche Phänomene, Konzepte und Repräsentationen aktiv miteinander verknüpfen (Opfermann et al., 2017; Treagust, 2018; Yore & Hand, 2010).

Bezogen auf die fachlichen Schwierigkeiten der Schüler:innen mit Repräsentationen umfasst das Fachgebiet der Chemie allerdings zahlreiche abstrakte und komplexe Konzepte (z. B. Oxidationszahlen), die wiederum mit vielen unterschiedlichen und schwierigen Repräsentationen wiedergegeben werden (Permatasari et al., 2022). Viele Lernende sind hingegen nur begrenzt in der Lage, abstrakt genug zu denken, was dazu führt, dass chemische Konzepte nur schwer verstanden werden und falsche Vorstellungen entstehen (Milenković et al., 2014). Schüler:innen der Chemie fällt es grundsätzlich schwer, einzelne Repräsentationen zu interpretieren, insbesondere wenn diese abstrakte Konzepte wiedergeben. Dies ist vor allem der Fall, wenn Schüler:innen wenig Erfahrung mit dem Arbeiten mit fachtypischen Repräsentationen haben und anhand dieser chemischen Inhalte lernen sollen. Beispielsweise können sie Schwierigkeiten haben, wenn sie chemische Reaktionen anhand einer chemischen Reaktionsgleichung verstehen oder den Aufbau eines Atoms anhand eines Atom-Schalen-Modells interpretieren sollen. Zusätzlich sind Schüler:innen nur

bedingt in der Lage, chemische Inhalte in Repräsentationen den Ebenen zuzuordnen. Sie können daher schwer unterscheiden, ob eine Repräsentation ein Abbild eines chemischen Phänomens ist oder ein Modell zur Erklärung der Struktur ist, welches nicht der Realität entspricht. Dies kann zu Fehlvorstellungen führen, wie z. B. der Annahme, dass Wassermoleküle tatsächlich wie die bekannte Strukturformel von Wasser aussehen, wenn man die Materie des Wassers stark genug vergrößern würde. Diese Schwierigkeiten mit Repräsentationen und die Entwicklung potentieller Fehlvorstellungen multiplizieren sich, wenn mehrere Formen und Ebenen von Repräsentationen zugleich eingesetzt werden. So besteht die Gefahr, dass falsch interpretierte Informationen aus einer Repräsentation in eine weitere überführt werden und falsche Informationen miteinander kombiniert und ergänzt werden (Permatasari et al., 2022).

Empirische Studien belegen zugleich, dass multiple externe Repräsentationen das Lernen chemischer Konzepte erleichtern können (z. B. Bodner & Domin, 2000; Widarti et al., 2019; Wiyarsi et al., 2018). Der positive Effekt multipler Repräsentationen wurde in Bezug auf chemische Themen untersucht und belegt, wie beispielsweise für Stoffumwandlungen und das chemische Gleichgewicht (Derman & Ebenezer, 2020; Priyasmika, 2021; Tima & Sutrisno, 2018). Dabei konnten nicht nur kognitive Wissenszuwächse festgestellt werden, sondern auch die Erhöhung der Selbstwirksamkeitserwartung bezüglich des Chemielernens sowie die Reduktion von Fehlvorstellungen und des Cognitive Loads (Permatasari et al., 2022; Tima & Sutrisno, 2018; Widarti et al., 2021).

Die letztliche Wissenskonstruktion chemischer Phänomene findet erst statt, wenn beim Lernen mit multiplen Repräsentationen auch die chemischen Repräsentationsebenen berücksichtigt werden. Dies erfolgt, wenn chemische Prozesse und Phänomene auf unterschiedlichen Repräsentationsebenen behandelt und diese anschließend aus unterschiedlichen Perspektiven miteinander verknüpft werden. Dieser Verknüpfungsprozess ist allerdings für viele Lernende anspruchsvoll und fehleranfällig, weil Lernende bereits große Schwierigkeiten mit den einzelnen Repräsentationsebenen aufweisen (Cheng & Gilbert, 2009; Gilbert & Treagust, 2009a; Treagust & Chandrasegaran, 2009). Eine zu geringe praktische bzw. experimentelle Erfahrung von Lernenden führt beispielsweise zu einem geringen makroskopischen Wissen. Gerade für Novizen ist dieses Wissen unentbehrlich, weil die makroskopische Ebene ihr erster Zugang zur Chemie

ist (Gilbert & Treagust, 2009b). Auf den anderen beiden Ebenen kann zudem eine fehlende Modellkritik zu falschen Vorstellungen führen. Generelle Limitationen von Repräsentationen auf der submikroskopischen und symbolischen Ebene sollten den Lernenden daher bekannt sein. So sollte klar sein, dass Wassermoleküle in der Realität weder eine blaue Farbe haben noch wie ein eingekreistes „H₂O“ aussehen (Scheiter & Richter, 2015). Ein weiteres Problem für Lernende sind die Abkürzungen und Symbole in der symbolischen Ebene, die aus Trivialnamen, Strukturformeln und der Nomenklatur bekannt sind. Als teilweise hochentwickelte und abstrakte Systeme gleichen sie einer eigenen Sprache und sind Novizen nur schwer zugänglich (Cheng & Gilbert, 2009; Sliwka, 2003; Taber, 2009).

Da diese und weitere ähnliche Schwierigkeiten bei den Schüler:innen auch über die Sekundarstufe I hinaus bestehen (Magnus et al., 2020), ist es für Lehrkräfte von Bedeutung, entsprechende Berücksichtigungen und Maßnahmen dauerhaft und auch bei erfahrenen Schüler:innen der Chemie vorzunehmen.

2.1.4 Kognitionspsychologische und fachdidaktische Gestaltungsprinzipien

Die dargelegten Lerntheorien und Eigenschaften von Repräsentationen liefern Implikationen, die bei der Gestaltung multipler externer Repräsentationen grundlegend berücksichtigt werden sollten (z. B. Gilbert & Treagust, 2009b; Mayer, 2014a; Treagust & Tsui, 2013b). Während sich die Kognitionspsychologie vor allem auf allgemeine Gestaltungsmerkmale konzentriert, die die Informationsverarbeitung erleichtern sollen (Ainsworth, 2006; Mayer, 2009c, 2009d), befasst sich die Chemiedidaktik vor allem mit inhaltlichen Aspekten, wie der Darstellung von Informationen auf der makroskopischen, symbolischen und submikroskopischen Ebene (Talanquer, 2011). Multiple externe Repräsentationen sollten grundsätzlich so gestaltet werden, dass sie während des Lernprozesses Informationen klar und verständlich wiedergeben und eine aktive Auseinandersetzung mit dem Lernmaterial einfordern. Gleichzeitig sollte dabei eine möglichst geringe kognitive Belastung ausgelöst werden (Schnotz, 2014). Im Folgenden werden die Anforderungen an multiple externe Repräsentationen auf Grundlage der benannten Implikationen und Gestaltungsprinzipien in fünf Kategorien beschrieben (nach Tonyali et al., 2023). Es ist dabei zu berücksichtigen, dass das Planen und Vorbereiten multipler Repräsentationen bzw. Lehr-

Lern-Materialien in der Unterrichtspraxis in der Regel so erfolgt, dass Lehrkräfte bereits bestehende Repräsentationen und Materialien nutzen, diese anpassen und einsetzen. So wird davon ausgegangen, dass Lehrkräfte bei der Materialsuche mit einer analytischen und selektiven Haltung vorgehen, fertige Materialien auf Basis ihrer fachlichen und fachdidaktischen Kompetenzen prüfen und sie an die Bedürfnisse ihrer Lerngruppe und -situation anpassen (Adamina, 2014). Nach diesem analytischen und prüfenden bzw. überarbeitenden Ansatz sind die folgenden fünf Kategorien zusammengefasst und strukturiert, sodass sie von einer Lehrkraft durchlaufen werden können, wenn sie eine multiple externe Repräsentation mit einem festgelegten Lernziel für eine spezifische Lerngruppe gestaltet. Es wird daher davon ausgegangen, dass der Lehrkraft dabei ein grundlegendes Material bereits vorliegt, welches von ihr anhand der Kategorien optimiert wird.

Kategorie 1: Passung der Repräsentationen zum Lernziel und Eignung für die Lerngruppe

Die erste Kategorie behandelt allgemeine Aspekte einer multiplen externen Repräsentation. Im Idealfall enthält eine multiple externe Repräsentation nur diejenigen Informationen, die für das Erreichen des Lernziels ausreichend sind. Das Unterrichtsmaterial ist prägnant, damit sich die Lernenden auf den Inhalt konzentrieren können und nicht durch irrelevante Informationen abgelenkt werden (Harp & Mayer, 1998). Daher ist es im ersten Schritt wichtig, festzulegen, welche chemischen Inhalte in ein Lehr-Lern-Material aufgenommen werden und welche nicht. Ist das Lernziel beispielsweise die Einführung in das Teilchenmodell, so enthält das Arbeitsblatt nur Informationen, die sich auf dieses Thema beziehen. Themen, die über das Lernziel hinausgehen, wie z. B. das Verhalten der Teilchen bei höheren Temperaturen, werden auf diesem Arbeitsblatt noch nicht behandelt.

Sobald die relevanten Inhalte identifiziert wurden, wird im nächsten Schritt überlegt, wie die Informationen zwischen Text und Bild verteilt werden und wie beide Darstellungen gestaltet sind. Hierbei sollte berücksichtigt werden dass, Theorien und allgemeingültige oder abstrakte Aussagen in den textlichen Teil der multiplen externen Repräsentation gehören, während Beispiele und konkrete

Anwendungen in Bildern gezeigt werden (Schnotz, 2002b; Schnotz & Bannert, 1999).

Kategorie 2: Gestaltung der bildbasierten Repräsentationen

Neben der Betrachtung der multiplen externen Repräsentation als Ganzes ist es notwendig, die verschiedenen Teile als einzelne Repräsentationen zu betrachten. Zunächst werden die Bildtypen identifiziert, um zu prüfen, ob ihre charakteristischen Eigenschaften korrekt umgesetzt wurden. So werden beispielsweise in der Chemie definierte Konventionen für Diagramme mit Koordinatensystemen eingehalten. Weitere und detailliertere Überlegungen betreffen auch Bilder, in denen zum Beispiel Moleküle im Teilchenmodell als Kugeln dargestellt werden. Die Farben und Formen der einzelnen Molekülarten sowie die Größenverhältnisse zueinander können das konstruierte Wissen beeinflussen und ggf. zu falschen Vorstellungen führen (Beispiel: „Wassermoleküle sind rund und blau.“). Im Hinblick auf die Repräsentationsebenen ist es notwendig, die Position der Elemente innerhalb einer bildlichen Repräsentation zu analysieren (Mayer, 2014b; Tan et al., 2009). Beispielsweise liegen Elemente einer Abbildung, die zur gleichen Ebene gehören (z. B. Zuckerteilchen und Wasserteilchen auf submikroskopischer Ebene), räumlich nahe beieinander. Elemente, die wiederum nicht zusammengehören, sind räumlich getrennt voneinander (z. B. ein Zuckerwürfel auf makroskopischer Ebene und Zuckermoleküle auf submikroskopischer Ebene). Ein weiterer, zu berücksichtigender Aspekt ist die Hervorhebung von Elementen in einem Bild. Das Hervorheben kann den Schüler:innen dabei helfen, das Bild besser zu verstehen, weil sie die wichtigsten Merkmale direkt erkennen (Mayer, 2009b).

Kategorie 3: Gestaltung der textbasierten Repräsentationen

Wie die bildlichen Repräsentationen lassen sich auch die deskriptiven Repräsentationen in verschiedene Kategorien und Typen einteilen. Grundsätzlich kann zwischen Basis- und Zusatztexten unterschieden werden. Basistexte liefern alle Informationen, die für das Erreichen des Lernziels wichtig sind. Im Gegensatz dazu kann ein Zusatztext optional als Zusammenfassung, mit Zusatzinformationen oder begleitenden Aufgaben hinzugefügt werden, um die Verständlichkeit des Basistextes zu erleichtern (Ballstaedt, 1997). Basis- und Zusatztext in einem Lehr-Lern-Material sind so gestaltet, dass die Lernenden

wichtige Informationen (Basistext) von eher unwichtigen Informationen (Zusatztext) auch optisch durch die Anordnung und Formatierung des Gesamttextes unterscheiden können (Ude, 2016).

Sobald der sachbezogene Text, bestehend aus Basis- und Zusatztext, sortiert und angeordnet ist, eignet es sich, den bereits bestehenden Text mit einem Kontext (Barke & Harsch, 2011; Bennett et al., 2006) zu ergänzen, welcher ein Problem, eine Frage oder eine Aufgabe aufwirft und den Lernenden einen Lernanlass bietet. Anschließend wird der gesamte Text auf der Absatz-, Satz- und Wortebene analysiert. Die Informationen werden in richtiger logischer und chronologischer Reihenfolge sowie in adäquater Fachsprache dargeboten. Jedes Wort im Text wird durchdacht verwendet, ist präzise und eindeutig und bietet keinen Raum für Missverständnisse. Ebenso ist es wichtig, Fachbegriffe zu definieren, wenn sie zum ersten Mal verwendet werden, und sie einheitlich zu verwenden. Sind zudem Schlüsselbegriffe typografisch hervorgehoben (fett, kursiv, farbig etc.), können die wichtigsten Informationen beim ersten Überfliegen eines Textes direkt erfasst werden (Ballstaedt, 1997). Eine solche Umformatierung des Textes kann die sensorische Aufnahme und kognitive Verarbeitung des Textes vereinfachen. Je einfacher die Schriftart und je schmaler die Textspaltenbreite ist, desto einfacher erfolgt der sensorische Leseprozess. Daher bietet es sich an, den Basistext wie in Zeitungsartikeln in Spalten zu unterteilen (Ballstaedt, 1997).

Kategorie 4: Gestaltung der multiplen externen Repräsentation

Nachdem die einzelnen Repräsentationen hinsichtlich ihrer Qualität optimiert wurden, kann im nächsten Schritt ihr Zusammenspiel als multiple externe Repräsentation untersucht und überarbeitet werden. Dabei ist es notwendig, die Funktion der Kombination bzw. des Zusammenwirkens zu berücksichtigen (Ergänzen, Einschränken, Vertiefung; s. Abbildung 2; Ainsworth, 1999). Auf diese Weise lässt sich auch bestimmen, welche Rolle jede einzelne Repräsentation innerhalb der ganzheitlichen multiplen externen Repräsentation übernimmt. Trägt beispielsweise die Abbildung die Hauptinformation und unterstützt sie die Informationen im Text oder sind die Funktionen umgekehrt verteilt? Sind die Rollen der einzelnen Repräsentationen identifiziert, können die Repräsentationen

besser aufeinander abgestimmt und miteinander verknüpft werden (z. B. „In der Abbildung 1 wird dieser Sachverhalt noch einmal visuell veranschaulicht.“).

Aus chemischer Sicht ist es außerdem wichtig, die dargebotenen Repräsentationsebenen und die gezeigten Beziehungen zwischen ihnen zu analysieren. Die Ebenen sind korrekt abgebildet und werden nicht vermischt. Dies gilt sowohl für den textlichen als auch bildlichen Teil einer multiplen externen Repräsentation. Beispielsweise ist es ungeeignet, Wassermoleküle (submikroskopisch) in einem Wasserglas (makroskopisch) abzubilden. In diesem Fall können falsche Vorstellungen hervorgerufen werden (z. B. Gilbert & Treagust, 2009a; Kozma & Russell, 1997; Treagust & Tsui, 2013a).

Kategorie 5: Auswahl und Formulierung von Aufgabenstellungen

Die aktive und tiefgreifende Auseinandersetzung der Lernenden mit den Elementen einer multiplen externen Repräsentation gilt als Voraussetzung für einen erfolgreichen Lernzuwachs. Diese kann nicht gelingen, ohne den Lernenden sorgfältig durchdachte und transparente Aufgabenstellungen vorzugeben (z. B. Draxler, 2005; Müller et al., 2002; Schmiemann, 2013). Als Hilfsmittel können hier zum Beispiel die curricular vorgegebenen Operatoren und Anforderungsbereiche des MSB NRW (2015) eingesetzt werden.

2.2 Repräsentationsbezogene Kompetenz von Chemielehrkräften

Um, wie im vorigen Kapitel 2.1 dargestellt, multiple externe Repräsentationen und Lehr-Lern-Materialien adaptieren und entwerfen zu können, benötigen angehende Lehrkräfte spezifische Fähigkeiten. Diese werden unter dem Begriff der repräsentationsbezogenen Kompetenz zusammengefasst (Kapitel 2.2.1) und lassen sich in die professionelle Kompetenz von Lehrkräften einordnen (Kapitel 2.2.2).

2.2.1 *Strukturierung der repräsentationsbezogenen Kompetenz*

In der Literatur wird die repräsentationsbezogene Kompetenz von Lehrkräften beispielsweise als „representational competence“ (Kozma & Russell, 1997, 2005) und „meta-representational competence“ (diSessa, 2004) beschrieben (s. Tabelle 1). In ihrer Arbeit integrierte Nitz (2012) beide Begrifflichkeiten und fasste ihre Konzepte in sechs Kategorien zusammen.

Die Definitionen von Kozma und Russel (2005) sowie von diSessa (2004) beschreiben Kombinationen von Fähigkeiten im Umgang mit Repräsentationen, die sich auf alle naturwissenschaftlichen Fächer übertragen und anwenden lassen. Während das Konzept „Representational competence“ (Kozma & Russell, 2005) domänenspezifische bzw. chemiebezogene Fähigkeiten beschreibt, ist das Konzept „Meta-representational competence“ (diSessa, 2004) eher domänenübergreifend und umfasst kreative Prozesse. Obwohl beide Konzeptionen den Umgang mit Repräsentationen aus unterschiedlichen Perspektiven betrachten, lassen sie sich in sechs Kategorien zusammenführen und als repräsentationsbezogene Kompetenz (übersetzt aus Nitz, 2012) benennen. Hierbei ist zu beachten, dass die Kategorien keine Hierarchie bilden, sondern simultan aufgebaut werden können (Nitz, 2012).

Tabelle 1: Definition repräsentationsbezogener Kompetenzen von Lehrkräften (verändert nach Nitz, 2012)

Representational competence (Kozma & Russell, 1997, 2005)	Kategorie (Nitz, 2012)	Meta-representational Competence (diSessa, 2004)
Merkmale und Eigenschaften von Repräsentationen identi- fizieren und beschreiben; Kon- zepte anhand von Repräsentationen beschreiben	Interpretation	Repräsentationen erläu- tern und erklären
Angemessene Repräsentationen auswählen und konstruieren	Konstruktion	Entwicklung und Design neuer Repräsentationen
Verschiedene Repräsentationen miteinander verbinden, ineinan- der überführen und die Bezie- hung zwischen ihnen erklären	Translation	Kritischer Vergleich ver- schiedener Repräsentati- onsformen und ihrer Angemessenheit
Angemessenheit und Zweckmä- ßigkeit verschiedener Repräsen- tationen erklären, ihre Aussagekraft beschreiben und vergleichen	Vergleich und Kritik	Kritischer Vergleich ver- schiedener Repräsentati- onsformen und ihrer Angemessenheit
Verstehen, dass Repräsentatio- nen mit Konzepten korrespon- dieren, aber auch Grenzen haben	Epistemo- logie und Funktions- weise	Verständnis des Zwecks von Repräsentationen ge- nerell und in bestimmten Kontexten, sowie der Art und Weise, wie Repräsen- tationen ihre Funktion er- füllen
Merkmale von Repräsentatio- nen nutzen, um im Diskurs Be- hauptungen zu unterstützen, Schlussfolgerungen zu ziehen und Vorhersagen über Relatio- nen zu treffen	Argumen- tation	—

2.2.2 Einordnung in die professionelle Handlungskompetenz

Die repräsentationsbezogene Kompetenz von Lehrkräften ist in hohem Maße domänenspezifisch (Kozma & Russell, 2005; Nitz, 2012). Sie kann in übergeordnete Kompetenzmodelle integriert werden, welche die gesamten Fertigkeiten einer Lehrkraft beschreiben, um die schulischen und unterrichtlichen Anforderungen zu bewältigen. Im deutschen Raum findet hierfür das Modell der professionellen Handlungskompetenz von Lehrkräften von Baumert und Kunter (2011) besondere Anwendung. Die Handlungskompetenz umfasst dabei nicht nur intellektuelle Fähigkeiten, sondern auch soziale, motivationale, intentionale und selbstregulative Fertigkeiten. Folgende Aspekte professioneller Kompetenz werden definiert: Professionswissen, Überzeugungen/Werthaltungen/Ziele, motivationale Orientierung und selbstregulative Fähigkeiten (Baumert & Kunter, 2011).

Die repräsentationsbezogene Kompetenz von Lehrkräften ist Teil des Professionswissens (Kleickmann et al., 2014). Das Professionswissen wird in fünf Kompetenzbereiche ausdifferenziert (Baumert & Kunter, 2011). Das *Fachwissen* (auch: content knowledge, CK) und das *fachdidaktische Wissen* (auch: pedagogical content knowledge, PCK) stellen den Kern der professionellen Kompetenz dar, da die Unterrichtsfächer den primären Handlungsrahmen einer Lehrperson bilden (Shulman, 1987). Neben diesem fachbezogenen Wissen werden drei fachunabhängige Kompetenzbereiche unterschieden: pädagogisch-psychologisches Wissen, Organisationswissen und Beratungswissen (Baumert & Kunter, 2011).

Das (allgemeine¹) Fachwissen und fachdidaktische Wissen werden als zwei getrennte Wissenskonstrukte betrachtet, wobei das fachdidaktische Wissen ein grundlegendes Fachwissen bedingt (Baumert & Kunter, 2011). Dementsprechend können bei Chemielehrkräften zwischen den beiden Konstrukten nicht nur schwache Korrelationen ($r_s = .323$, $p = .001$), sondern auch prädiktive Effekte festgestellt werden (Dollny, 2011). Die Ergebnisse gelten auch für angehende Lehrkräfte der Chemie, die sich im Vorbereitungsdienst befinden (Mutke, 2017). In Bezug auf die repräsentationsbezogene Kompetenz von Lehrkräften

¹ Zur besseren Unterscheidung von *repräsentationsbezogenem* Fachwissen und *repräsentationsbezogenem* fachdidaktischen Wissen wird das von Baumert und Kunter (2011) beschriebene übergeordnete und umfassende Fachwissen und fachdidaktische Wissen im Folgenden als *allgemeines* Fachwissen und *allgemeines* fachdidaktisches Wissen präzisiert.

greifen das allgemeine Fachwissen und das allgemeine fachdidaktische Wissen u. a. Inhaltsbereiche auf, die chemische Repräsentationen fokussieren (Baumert & Kunter, 2011; Kleickmann et al., 2014; Kleickmann et al., 2013). Diese Wissensbereiche werden im Folgenden als repräsentationsbezogenes Fachwissen sowie repräsentationsbezogenes fachdidaktisches Wissen definiert. Werden die Konzepte des allgemeinen Fachwissens und fachdidaktischen Wissens angelehnt an Kleickmann et al. (2014) auf die repräsentationsbezogenen Wissensbereiche übertragen, so lassen sich diese wie folgt präzisieren und definieren: Das *repräsentationsbezogene Fachwissen* fokussiert schulbezogene Inhalte und umfasst die sachbezogene Ebene von Repräsentationen. Diese Facette findet sich auch in den integrierten Kategorien von Nitz (2012) wieder (Interpretation, Konstruktion, Translation, Vergleich und Kritik). Das *repräsentationsbezogene fachdidaktische Wissen* geht über die Merkmale und Strukturen verschiedener Repräsentationen, Repräsentationsformen und -ebenen hinaus und umfasst auf sie bezogenes Wissen über Schülerkognitionen, Lernschwierigkeiten und Instruktionsstrategien. Die entsprechenden Ansätze lassen sich bereits in den Kategorien Vergleich und Kritik, Epistemologie und Funktionsweise sowie Argumentation wiederfinden und noch weiter auf das fachdidaktische Wissen präzisieren (nach Nitz, 2012). Die Kategorie „Vergleich und Kritik“ umfasst dabei Anteile sowohl repräsentationsbezogenen Fachwissens (z. B. Limitationen von Repräsentationsformen, Vergleich von Vor- und Nachteilen unterschiedlicher Repräsentationsformen) als auch fachdidaktischen Wissens (z. B. Auswahl von Lehrmethoden für Vergleich und Kritik, Vergleich der Eignung von unterschiedlichen Repräsentationsformen für eine Lerngruppe,). Inhaltlich gesehen, lässt sich demnach ein gradueller Übergang vom repräsentationsbezogenen Fachwissen in das fachdidaktische Wissen feststellen, obwohl beide Wissensbereiche zwei getrennte Konstrukte abbilden (z. B. Kleickmann et al., 2014; Kunter et al., 2011).

Die professionelle Handlungskompetenz nach Baumert und Kunter (2011) wurde bisher vor allem auf der Wissensebene betrachtet. Es zeigt sich jedoch, dass auch andere Aspekte in Bezug auf das Designen von multiplen externen Repräsentationen und Lehr-Lern-Materialien eine entscheidende Rolle spielen. Ein gefestigtes repräsentationsbezogenes Fachwissen und fachdidaktisches Wis-

sen reichen nicht aus, um diese Fertigkeiten auch unterrichtspraktisch umzusetzen. In diesem Zusammenhang sind die *Überzeugungen (Beliefs)* von Lehrkräften von Bedeutung. Diese sind subjektiv und beinhalten normative, fachbezogene Einstellungen zum Lehren und Lernen (Oser & Blömeke, 2012). Im Gegensatz zum Wissen gelten Überzeugungen jedoch als schwer veränderbar, da sie über lange Zeiträume aufgebaut werden (Pajares, 1992). Dennoch haben sie einen großen Einfluss auf die Unterrichtsgestaltung und die Unterrichtsqualität. Überzeugungen können als wichtige Prädiktoren und Mediatoren fungieren, die das Lehrverhalten und die Entscheidungen einer Lehrkraft maßgeblich prägen (Dubberke et al., 2008; Nitz, 2012). Adaptiert nach den fachsprachlichen Überzeugungen von Lehrkräften von Nitz (2012), sind die *repräsentationsbezogenen Überzeugungen* der Lehrkräfte von Bedeutung, die sich auf die Relevanz externer Repräsentationen für das Lernen von Chemie beziehen. Demnach setzen Lehrkräfte, die eine hohe Wertschätzung für Repräsentationen, Modellen oder Symbolen im Fach Chemie aufweisen, gezielt externe Repräsentationen ein, um das Lernen der Schüler:innen zu fördern.

Ein weiterer relevanter Aspekt für das unterrichtspraktische Handeln ist die *Motivation* einer Lehrkraft, die eng mit der *Selbstwirksamkeitserwartung* verknüpft ist. Sie beeinflusst das Engagement, die Resilienz und emotionale Wahrnehmung im schulischen Kontext. Ebenso relevant ist die *Selbstregulation* der Lehrkraft bzw. ihr Verhalten im Hinblick auf gesetzte Ziele. Bezogen auf die generelle pädagogische Handlungskompetenz haben diese Faktoren einen maßgeblichen Einfluss auf die Entwicklungsbereitschaft von Lehrkräften (Kunter et al., 2011; Kunter et al., 2013). Lehrkräfte mit hoher Motivation und Selbstwirksamkeitserwartung zeigen eine stärkere Entwicklungsbereitschaft und sind offener für kritisches Feedback und für den Austausch mit Kolleg:innen. Die Offenheit für Feedback fördert wiederum die kontinuierliche professionelle Weiterentwicklung (Hattie & Timperley, 2007), sodass sich beide Faktoren gegenseitig positiv beeinflussen. Selbstreguliertes Verhalten und ausgeprägte Motivation ermöglichen Lehrkräften daher, ihre beruflichen Herausforderungen aktiv zu steuern und die eigene Praxis zu hinterfragen, was zu einer höheren Unterrichtseffektivität führen kann (Kunter et al., 2011).

2.2.3 Entwicklung der Kompetenz im Rahmen der Lehrkräftebildung

In den Beschlüssen der Kultusministerkonferenz werden nicht nur die Bildungsstandards für Schüler:innen (z. B. KMK, 2004), sondern auch die Anforderungen an die Lehrkräftebildung festgelegt. Diese Regelungen umfassen sowohl fächerübergreifende (KMK, 2004) als auch fach- (KMK, 2008) und schulformspezifische Anforderungen (KMK, 1997a; 1997b). Darüber hinaus wird geregelt, in welcher Phase der Lehrkräfteausbildung welche Kompetenzen von den angehenden Lehrkräften erreicht werden sollen (KMK, 2012). Die Regelungen werden auf der Länderebene umgesetzt (z. B. MSB NRW, 2021; 2022).

Nach dem Abschluss der ersten Phase der Lehrkräftebildung mit einem universitären Hochschulabschluss absolvieren die angehenden Lehrkräfte einen in der Regel 18-monatigen Vorbereitungsdienst. Besonders diese zweite Phase der Lehrkräftebildung stellt für die angehenden Lehrkräfte eine Herausforderung dar, da in dieser Phase das im Studium erworbene theoretische Wissen auf praktische Anwendungssituationen übertragen werden muss (Kagan, 1992; Keller-Schneider & Hericks, 2011). Zahlreiche Studien weisen auf die Schwierigkeiten angehender Lehrkräfte im Umgang mit Repräsentationen hin, welche sowohl das repräsentationsbezogene Fachwissen als auch das fachdidaktische Wissen betreffen. Erstens haben angehende Lehrkräfte ein begrenztes Verständnis davon, wie Fachwissen repräsentiert und dargestellt wird. Dies betrifft auch den Umgang mit multiplen Repräsentationen (Cheng & Gilbert, 2009; Gilbert & Treagust, 2009c; Tsui & Treagust, 2013; van Driel & Verloop, 1999). In der Folge können Lehrkräfte die Qualität von Repräsentationen nur eingeschränkt beurteilen und bewerten. Zweitens verfügen Lehrkräfte nicht über ausreichende Kompetenzen im Umgang mit repräsentationsbezogenen Defiziten auf Schüler:ebene. Dadurch reflektieren sie den Umgang ihrer Schüler:innen mit Repräsentationen nicht und betrachten die Kommunikation über Repräsentationen als selbstverständlich und unproblematisch (McElvany et al., 2012). Infolgedessen werden Repräsentationen im naturwissenschaftlichen Unterricht lediglich als (fach-)didaktisches Hilfsmittel verwendet, jedoch nicht als eigenständiger Unterrichtsgegenstand (Taskin et al., 2015a). Tatsächlich stellen naturwissenschaftliche – insbesondere chemische – Repräsentationen für Schüler:innen eine erhebliche Herausforderung dar (Taber, 2009; Taskin & Bernholt, 2012).

Für den Aufbau repräsentationsbezogener Kompetenzen stehen angehenden Lehrkräften unterschiedliche Lerngelegenheiten zur Verfügung (Kleickmann & Anders, 2011; Klusmann, 2011; Richter, 2011). Es gibt zahlreiche Arbeiten (z. B. Dollny, 2011; Mutke, 2017; Strübe, 2021), die die Professionalisierung während des Vorbereitungsdiensts erforschen. Konkret wurde die repräsentationsbezogene Kompetenz jedoch kaum als Untersuchungsgegenstand in Betracht gezogen oder nur bei Lehramtsstudierenden untersucht (McElvany et al., 2009). Es wird daher angenommen, dass Lehrkräfte ihre Repräsentationskompetenz bezüglich des Vorbereitens und Designens von Repräsentationen im Rahmen der schulischen Berufsausübung selbstständig und individuell erwerben (McElvany et al., 2010). Dabei besteht jedoch die Gefahr, dass angehende Lehrkräfte Muster und Verhaltensweisen verinnerlichen, die möglicherweise nur ihrer Intuition entspringen und aus (fach-)didaktischer Sicht unbegründet, unreflektiert und eher lernhinderlich sind, obwohl die Lehrkräfte eigentlich das Gegenteil beabsichtigen.

Zur ersten Phase der Lehrkräftebildung liegen Befunde zum Aufbau repräsentationsbezogener Kompetenz vor: In einer Studie von Taskin et al. (2015a, 2015b) wurde festgestellt, dass das repräsentationsbezogene Fachwissen und fachdidaktische Wissen von Lehramtsstudierenden nicht auf das Studium selbst, sondern hauptsächlich auf die eigene Schulzeit zurückzuführen sind und während des Studiums wenig zunehmen. Darüber hinaus wurde festgestellt, dass nur Studierende mit einem hohen repräsentationsbezogenen Vorwissen in der Lage sind, ihr Wissen in einem ausreichenden Maß zu steigern. Studierende mit einem eher geringen Vorwissen zeigen hingegen kaum einen Wissenszuwachs während des Studiums. Außerdem wurde festgestellt, dass externe Repräsentationen selbst kaum als Lerninhalt behandelt werden (vgl. die Kategorien der repräsentationsbezogenen Kompetenz nach Nitz, 2012), sondern als Werkzeuge der Wissensvermittlung eingesetzt werden (Taskin et al., 2015b). Demnach gestalten Chemielehrkräfte im Berufseinstieg ihre Unterrichtsmaterialien bzw. externen Repräsentationen gemäß den Erfahrungen, die sie selbst als Schüler:innen gemacht haben (nach Hascher, 2014). Diese Erkenntnisse implizieren nicht nur die Notwendigkeit der Förderung des repräsentationsbezogenen Wissens angehender Lehrkräfte, sondern auch den Bedarf eines Anwendungsbezugs auf konkrete unterrichtspraktische Situationen, wofür sich der Vorbereitungsdienst anbietet.

So können die einzelnen Aspekte und Wissensbereiche der professionellen Handlungskompetenz gefördert werden. Zudem erhalten die angehenden Lehrkräfte durch unterrichtsnahe Beispiele eine direkte Verknüpfung von Theorie und Praxis.

2.3 Feedback zur Unterrichtsplanung von Chemielehrkräften

In der pädagogischen Psychologie wird Feedback als eine „bewusste Rückmeldung von Informationen an eine Person zu ihrem vorherigen Verhalten“ (Müller & Ditton, 2014, S. 15) definiert und ihr in Bezug auf seine Förderlichkeit eine besondere Relevanz und Wirksamkeit zugesprochen. Feedback kann einer Person gegeben werden, damit beispielsweise zukünftige Handlungsmuster optimierter ablaufen. Effektives Feedback ist jedoch auch von einem hohen Komplexitätsgrad gekennzeichnet, da es multidimensional in unterschiedlichen Kommunikations- und Gestaltungsformen gegeben und erhalten werden kann (Ditton & Müller, 2014). So gibt es in der Literatur zahlreiche Modelle, die die Bedingungen und Wirkungen von Feedback in Lehr-Lernsituationen beschreiben (z. B. Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Narciss, 2014; Ryan & Deci, 2000).

In Bezug auf die Lehrkräftebildung spielt Feedback während des Vorbereitungsdiens eine entscheidende Rolle, um die professionelle Entwicklung angehender Lehrkräfte zu unterstützen. Durch Feedback können angehende Lehrkräfte Informationen über ihre Stärken und Schwächen erhalten, wodurch ihnen ermöglicht wird, sich weiterzuentwickeln (Hattie & Timperley, 2007). Dabei erhalten die angehenden Lehrkräfte Rückmeldungen insbesondere von ihren Fachleitungen sowohl zu ihren Unterrichtsplanungen (z. B. didaktische und methodische Überlegungen, vorbereitetes Lehr-Lern-Material) als auch zu ihrem konkreten unterrichtspraktischen Handeln während der Unterrichtsstunden (z. B. Gesprächsführung, Umsetzung der Planungen). Die meisten empirischen Studien zum Lehrkräfteberuf im Zusammenhang mit Feedback beziehen sich allerdings auf Abschnitte der Lehrkräftebildung, die vor dem Vorbereitungsdienst stattfinden. Da aber Bewertungs- und Benotungspraktiken eng mit Feedback in Verbindung stehen und gerade im Vorbereitungsdienst unter starker Kritik stehen (Behnke, 2016), werden aktuelle Feedbackpraktiken im Vorbereitungsdienst sowie deren Effektivität vermehrt untersucht. Strietholt und Terhart (2009) erfassten beispielsweise, dass Fachleitungen ihre Bewertungen vorrangig mithilfe von Beobachtungen und Befragungen in Form von Unterrichtsbesuchen- und entwürfen, Nachbesprechungen sowie der zweiten Staatsarbeit vornehmen. Hierbei verwenden Fachleitungen zwar Instrumente, welche ihre

Bewertungskriterien verdeutlichen, nutzen diese aber in einem stark uneinheitlichen und unsystematischen Umfang (Kärner et al., 2019; Strietholt & Terhart, 2009).

Befunde aus der Forschung zur Lehrkräfteprofessionalisierung weisen dennoch auf das Potenzial von Feedback in der Aus- und Weiterbildung im Hinblick auf das praktische Lehrhandeln hin (Lipowsky, 2009). Es bleibt daher die Frage, inwieweit das Feedback- und Bewertungssystem im Vorbereitungsdienst optimiert werden und dadurch die Lehrkräftebildung profitieren kann, beispielsweise durch den Einsatz prozessorientierten, kontinuierlichen und gezielten Feedbacks (Hascher, 2014). Doch zunächst gilt es dafür genauer zu betrachten, was solch ein korrigierendes bzw. formatives und tutorielles Feedback auszeichnet und welche Prozesse während des Generierens und Verarbeitens von Feedback ablaufen.

2.3.1 Informatives tutorielles Feedback

Die Ergebnisse zu empirischen Untersuchungen über effektives Feedback sind insgesamt inkonsistent und teilweise widersprüchlich. Diese Inkonsistenz deutet daraufhin, dass die Wirksamkeit von Feedback von situativen und individuellen Faktoren abhängt (Lipnevich & Panadero, 2021; Müller & Ditton, 2014). Um dennoch zentrale und relevante Faktoren von Feedback zu benennen, wird das ITFL-Modell („Interactive Two Feedback Loops Model“, Narciss, 2006) vorgestellt, in welchem die bisher entwickelten Feedbackmodelle von Kulhavy und Stock (1996), Bangert-Drowns und Kolleg:innen (1991), Butler und Winne (1995) sowie Kluger und DeNisi (1996) erweitert wurden. Das ITFL-Modell eignet sich für die Zielgruppe angehender Lehrkräfte besonders gut, da es nicht nur die individuellen Faktoren der Feedbackerhaltenden berücksichtigt, sondern auch die inhaltlichen und formalen Komponenten des Feedbacks. Darüber hinaus werden Prinzipien für die Gestaltung und Evaluation informativer und tutorieller Feedbackarten benannt (Narciss, 2014).

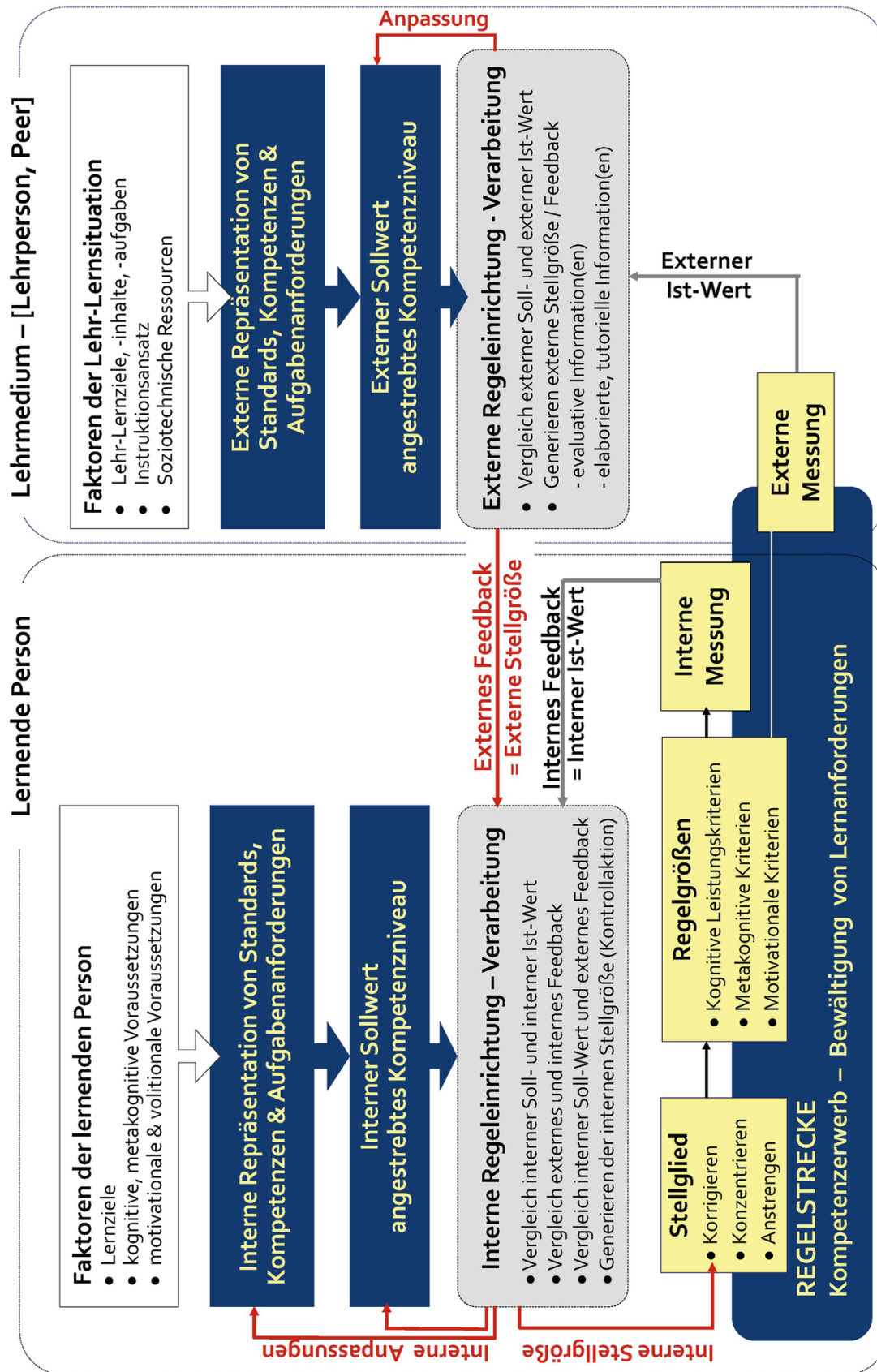


Abbildung 5: Interactive Two Feedback-Loops Model (ITFL-Modell; aus Narciss, 2018)

Informatives tutorielles Feedback wird als eine intervenierende Maßnahme gesehen mit dem Ziel, Lehr-Lernprozesse so zu unterstützen, dass Lernende Wissen und Fertigkeiten erwerben, um eine gegebene Aufgabe zu bewältigen oder Anforderung zu erreichen. Maßgebend für die Effektivität des Feedbacks ist die Qualität des erhaltenen Feedbacks, (z. B. Inhalt, Umfang, Art und Struktur), der individuellen Voraussetzungen des Lernenden (z. B. Vorwissen, metakognitive Strategien, Motivation) und situative Faktoren des Lehr-Lern-Kontexts (Lernziele, Lerninhalte, Aufgaben, Anforderungen; Narciss, 2014). Der Feedbackprozess selbst wird anhand einer Regelungstechnik veranschaulicht und besteht aus zwei Feedbackschleifen, welche jeweils intern und extern ablaufen (Abbildung 5). Um einen Feedbackprozess zu starten, ist es zunächst notwendig, die Regelgröße zu bestimmen. In einer Lehr-Lern-Situation ist dies die Kompetenz des Lernenden, die durch das Feedback reguliert werden soll. Der Inhalt der Regelgröße und der zu erbringende Umfang kann dabei durch Kriterien festgelegt werden. Der Ist-Wert dieser Regelgröße kann jeweils durch eine interne und externe Messung sowie anschließende Feedbackschleife evaluiert werden. Der interne Feedbackprozess erfolgt dabei durch die lernende Person selbst, während das externe Feedback durch ein Lehrmedium bzw. durch eine Lehrperson oder Peers generiert und an die lernende Person gegeben wird. Dabei bringen sowohl die lernende als auch die lehrende Person individuelle Faktoren bzw. Voraussetzungen und Feedbackfähigkeiten mit, welche sich ebenso in ihren individuellen Standards, Kompetenzen und Aufgabenanforderungen widerspiegeln. Feedback, welches sich selbst oder einer anderen Person gegeben wird, ist daher immer von subjektiven Einschätzungsmerkmalen geprägt. Aus der internen oder externen Messung wird dann ein interner bzw. externer Sollwert bzw. ein angestrebtes Kompetenzniveau entwickelt. Zentral für die Generierung des Feedbacks ist im nächsten Schritt die interne oder externe Regeleinrichtung, in der der angestrebte Sollwert mit der gemessenen Regelgröße bzw. dem Ist-Wert verglichen werden. In diesem Prozess wird das interne Feedback generiert und verarbeitet. Zudem wird hier auch das externe Feedback erhalten und gemeinsam mit dem internen Feedback verarbeitet, sodass im Stellglied der lernenden Person regulierende Maßnahmen ergriffen werden und die Regelgröße an den Soll-

Wert angepasst wird. Das Stellglied, die Regelgröße sowie die interne und externe Messung bilden dabei die Regelstrecke bzw. den Kompetenzerwerb (Narciss, 2018).

Damit das generierte Feedback möglichst effektiv reguliert wird bzw. der Ist-Wert der Regelgröße so gut wie möglich an das angestrebte Kompetenzniveau angepasst werden kann, gilt es drei Bedingungen zu beachten (Narciss, 2014):

- *Präzise Bestimmung der Anforderungen und Lehr-Lernziele:* Die zu optimierende Regelgröße sollte sowohl im Ganzen als auch in Teilen sorgfältig definiert und operationalisiert sein, ebenso wie die Regelgröße zu messen ist und mit welchen Maßnahmen sie korrigiert werden kann. In einer Lehr-Lern-Situation sollten daher Lehr-Lern-Ziele, Lerninhalte und zu bewältigende Aufgaben für die lernende Person transparent sein.
- *Qualitätsfaktoren der externen Feedbackschleife:* Auch hier erfordert es ein maximales Maß an Transparenz. Die jeweils internen und externen Soll-Werte und Kompetenzanforderungen sollten möglichst gleich aussehen. Des Weiteren sollte für das externe Feedback ein reliables und valides Diagnoseinstrument eingesetzt werden, welches sowohl den Ist-Wert erfasst als auch den Soll-Wert angibt und somit auch deren Diskrepanz. Wurden so die Feedbackinformationen festgelegt, muss zuletzt überlegt werden, wie diese an die lernende Person herangetragen werden, damit das externe Feedback so vollständig wie möglich von dem Lernenden verarbeitet werden kann.
- *Qualitätsfaktoren der internen Feedbackschleife:* Ziel eines wiederholenden internen Feedbacks ist die Befähigung der lernenden Person, sich selbstständig zu evaluieren und erhaltenes externes Feedback effektiv zu verarbeiten. Die Person soll dadurch lernen, optimierende Maßnahmen aus dem gesamten Verarbeitungsprozess zu erschließen und umzusetzen. Relevant ist auch hier, wie bei der externen Messung, wie die Ist-Werte des internen Feedbacks bestimmt werden, was mit einer hohen Selbsteinschätzungsfähigkeit einhergeht. Letztlich entscheiden die individuellen Faktoren bzw. maßgeblich die Motivation und das Interesse des Lernenden, ob und wie vollständig die Feedbackverarbeitungsprozesse in der internen Regeleinrichtung und Regelstrecke durchgeführt werden (Narciss, 2014).

Während sich das ITFL-Modell mehrheitlich auf die Bedingungen und das Format des Feedbacks fokussiert, bleibt die Frage offen, wie die inhaltliche Ebene des Feedbacks gestaltet werden soll. Hierzu liefert das Feedback-Modell nach Hattie und Timperley (2007) ergänzende Informationen. Nach diesem Modell beinhaltet effektives Feedback folgende Komponenten und kann folgende zugehörige Fragen aus der Perspektive des Lernenden beantworten:

- *Feed up*: Wo stehe ich und was ist das Ziel?
- *Feed back*: Welche Fortschritte werden auf dem Weg zum Ziel gemacht? Wie habe ich mich verbessert?
- *Feed forward*: Welche Maßnahmen müssen ergriffen werden, um bessere Fortschritte zu erzielen?

Somit ist ebenso eine zeitliche Komponente in das Feedback eingebunden, die den individuellen Entwicklungsverlauf des Lernenden mitbegutachtet und berücksichtigt (Hattie & Zierer, 2025).

Zusammenfassend gibt es eine große Varianz an Prozessmechanismen und Parametern, die Hinweise dafür liefern, wie Feedback gestaltet werden soll, damit es möglichst effektiv, korrigierend und formativ wirkt. Wie diese Hinweise praktisch und zur Professionalisierung angehender Lehrkräfte umgesetzt und in Arbeitsprozesse eingebunden werden können, wird in den Implikationen (Kapitel 2.3.2) beschrieben.

2.3.2 Implikationen zur Feedbackgestaltung für (angehende) Lehrkräfte

Die Bedingungen der Wirksamkeit von (tutoriellem) Feedback lassen sich sowohl auf entwickelte Feedbackmodelle als auch empirische Studien aus dem Bereich der Unterrichtspraxis zurückführen. Basierend auf dem ITFL-Modell lassen sich Implikationen herausstellen, die einerseits die Multidimensionalität sowie -funktionalität von tutoriellem Feedback betonen und andererseits Gestaltungsprinzipien vorschlagen. Wichtig ist vor allem die Passung zwischen dem Feedback und (1) den zu bewältigenden Aufgaben, (2) den gemachten Fehlern und (3) dem Wissensniveau der lernenden Person. Um diese Passungen herzustellen, müssen folgende Faktoren beachtet werden (Narciss, 2014):

- Funktion, Inhalt und Form des Feedbacks
- Merkmale des Lehr-Lernkontexts (Lehr-Lernziele, Lerninhalte, Aufgaben, Anforderungen)

- Individuelle Voraussetzungen und Merkmale der lernenden Person (Vorwissen und Fertigkeiten, Lernmotivation)

Tabelle 2: Klassifikation unterschiedlicher Feedbackformen (nach Narciss, 2006, S. 23)

Bezeichnung	Beispiele für Feedbackinhalte
<i>Einfache Formen</i>	
Ergebnis-/Antwortbezogenes Feedback (Knowledge of result/response)	Falsch/richtig; nicht richtig/stimmt
Leistungsbezogenes Feedback (Knowledge of performance)	15 von 20 Aufgaben richtig; 85% der Aufgaben richtig gelöst
Lösungsbezogenes Feedback (Knowledge of correct result)	Angabe der korrekten Lösung; Markierung der korrekten Antwort
<i>Elaborierte Formen</i>	
Aufgabenbezogenes Feedback (Knowledge on task constraints)	Hinweise auf Art der Aufgabe, Bearbeitungsregeln, Teilaufgaben und Aufgabenanforderungen
Konzeptbezogenes Feedback (Knowledge about concepts)	Hinweise auf Fachbegriffe; Beispiele für Begriffe; Hinweise zum Begriffskontext; Erklärungen zu Begriffen
Fehlerbezogenes Feedback (Knowledge about mistakes)	Anzahl der Fehler; Ort der Fehler; Art der Fehler; Ursache der Fehler
Lösungsprozessorientiertes Feedback (Knowledge on how to proceed)	Fehlerspezifische Korrekturhinweise; aufgabenspezifische Lösungshinweise, Hinweise zu Lösungsstrategien; Leitfragen; Lösungsbeispiele
Metakognitives Feedback (Knowledge on meta-cognition)	Hinweise auf metakognitive Strategien; metakognitive Leitfragen

Das Feedback sollte inhaltlich so aufgebaut und formal gestaltet sein, dass es seine Funktion zum Erreichen des Lernziels erfüllt und zugleich die individuellen Voraussetzungen des Lernenden berücksichtigt. Wenn beispielsweise das Ziel darin besteht, dass die lernende Person nicht nur ihre Fehler korrigiert, sondern auch lernt, Fehler zu erkennen und Korrekturstrategien zu entwickeln, sollte das Feedback sowohl eine korrigierende als auch metakognitive Funktion haben. Der Umfang dieser Informationen sollte so umfassend sein, dass die lernende Person bei ihren individuellen Voraussetzungen abgeholt und an die Kompetenzanforderungen herangeführt wird. Diese Faktoren sind auch entscheidend dafür, welche Feedbackform eingesetzt wird und welchen konkreten Inhalt das Feedbacks aufweist (Tabelle 2). Inhaltlich betrachtet werden zudem die evaluative und informative Komponente. Durch die evaluative Komponente wird der erreichte Leistungsstand erfasst, wobei durch die informative Komponente beispielsweise Hinweise zu zukünftigen Verbesserungsschritten angegeben werden (Narciss, 2014). Die formale Gestaltung des Feedbacks bezieht sich auf das Format, in dem das Feedback präsentiert wird. Das Feedback kann z. B. schriftlich oder mündlich gegeben werden, unmittelbar oder verzögert, graphisch oder sprachlich (Hattie & Wollenschläger, 2014). Wie die Parameter ausgewählt werden, hängt von der Feedbackfunktion und vom Lernziel des Feedbackprozesses ab (Narciss, 2014).

Ist das Feedback generiert, stellt sich im nächsten Schritt die Frage, wie diese Informationen präsentiert bzw. kommuniziert werden. Hier schlägt das ITFL-Modell folgenden Feedbackprozess vor:

1. *Feststellen der Regelgrößen und des Soll-Werts*: Es wird klar vorgegeben, welche Regelgrößen Gegenstand des Feedbacks sind und welche Kompetenzanforderungen gelten. So kann gewährleistet werden, dass die lernende Person eine möglichst passende interne Repräsentation (vgl. Abbildung 5) bildet.
2. *Anregung des internen Feedbacks*: Erst wird die interne und anschließend die externe Feedbackschleife angeregt, um den Schwerpunkt des Vorgehens beim selbstregulierten Lernen beizubehalten. Das selbstregulierte Lernen wird zunächst alleinständig aktiviert, bevor es vom externen Feedback beeinflusst wird.

3. *Darbieten externen Feedbacks*: Das externe Feedback soll Übereinstimmungen und auch Diskrepanzen zwischen dem Ist- und Soll-Wert des Lernenden verdeutlichen, die dem Lernenden durch das interne Feedback möglicherweise nicht ausreichend deutlich wurden.
4. *Anregung von Reflexion und Verbesserungsmaßnahmen*: Damit die Selbstregulation erfolgreich ablaufen kann, muss sich die lernende Person sowohl mit dem internen als auch externen Feedback auseinandersetzen und selbstständig Korrekturmaßnahmen vornehmen. Falls dies nicht selbstständig möglich ist, können Hinweise, Erklärungen oder Beispiele dargeboten werden.
5. *Darbietung einer Anwendung bzw. eines Transfers der Feedbackinformationen*: Nach dem Prinzip eines nachhaltigen Lernprozesses werden neu gelernte Fähigkeiten von Lernenden angewandt und erprobt. Dies gilt auch für die neu erlernten Erkenntnisse, die aus der Verarbeitung des internen und externen Feedbacks hervorgehen. Auf diese Weise werden die Lernerfolge und -fortschritte erkenntlich, sodass bei Bedarf weitere Feedbackschleifen und Optimierungsprozesse nachgeschaltet werden können (Narciss, 2014).

Darüber hinaus liegen Befunde und Reviews zur Feedbackforschung vor, die sich konkret auf die Feedbackgestaltung von Lehrkräften und für Lehrkräfte beziehen. Zentrale Hinweise liefert hierzu beispielsweise das Review von Scheeler et al. (2004), in dem beschrieben wird, dass für eine dauerhafte Optimierung:

- das Darbieten von Feedback wirksamer ist als kein Feedback darzubieten,
- unmittelbares Feedback wirksamer ist als verzögertes Feedback,
- unmittelbares, spezifisches, positiv formuliertes und korrigierendes Feedback am erfolgversprechendsten ist.

Weitere Kriterien bietet die Metaanalyse von Brinko (1993), die speziell die Bedarfe von Unterrichtenden bzw. Lehrkräften in den Blick nimmt. Hier werden die Bedingungen wirksamen Feedbacks nach den W-Fragewörtern kategorisiert und erklärt. Eine selbsterstellte Übersicht der zentralen Befunde ist in Tabelle 3 zu finden.

Als wichtig stellt sich heraus, dass Feedback aus unterschiedlichen Perspektiven bzw. Quellen (intern, extern, tutoriell, kollegial etc.) und verschiedenen Formaten erhalten werden. Gleichzeitig sollte das Feedback präzise, beschreibend und

unterstützend gestaltet sein, damit unterrichtspraktische Handlungen optimiert werden können. Relevant ist zudem, dass das Arbeiten mit Feedback nicht nur als eine punktuelle oder einmalige Korrektur aufgefasst wird, sondern als ein überdauernder kontinuierlicher Prozess (Brinko, 1993).

Ausgehend von den Befunden von Brinko (1993), Butler und Winne (1995) sowie Hattie und Timperley (2007) wird außerdem angenommen, dass vor allem die Kombination aus internem Feedback (Selbstfeedback) und externem Feedback (Fremdfeedback) geeignet ist, angehende Lehrkräfte zu fördern (Wollenschläger et al., 2012). Während das interne Feedback von der lernenden Person selbst generiert wird, kann das externe Feedback von unterschiedlichen Quellen gegeben werden. Externe feedbackgebende Instanzen im Kontext des Vorbereitungsdienst reichen von Vorgesetzten (z. B. Fachleitung, Schulleitung, Ausbildungslehrkraft) über Kolleg:innen (z. B. Fachkollegium, Mitreferendar:innen) bis hin zu Schüler:innen, welche das erbrachte unterrichtspraktische Handeln rückmelden (Müller & Ditton, 2014). Damit können auch Lehrkräfte und Lehramtsanwärter:innen externe Feedbackquellen sein, wenn sie zum Beispiel ihre Schüler:innen bewerten. Allerdings ist das Feedback, welches von angehenden Lehrkräften ausgeht, im Vergleich zum Feedback, welches an Lehramtsanwärter:innen gerichtet ist, empirisch betrachtet deutlich besser untersucht (Bitan, 2014).

Tabelle 3: Bedingungen der Wirksamkeit von Feedback zur Verbesserung des Unterrichtens (Eigene Zusammenstellung nach Brinko, 1993)

Parameter von Feedback und zugehörige Bedingungen
<i>Quelle des Feedbacks (Wer?)</i>
- Die Informationen werden aus verschiedenen Quellen gesammelt, sowohl von der empfangenden Person selbst als auch von anderen Personen. - Die Feedbackquelle wird als glaubwürdig, sachkundig, wohlmeinend, unterstützend, respektvoll und authentisch wahrgenommen. - Die Quelle des Feedbacks hat den gleichen oder einen anderen Status wie die empfangende Person des Feedbacks.

Empfangende Person des Feedbacks (Wer?)

- Die empfangende Person nimmt das Feedback freiwillig an und/oder nimmt den Feedbackprozess als Teil der routinemäßigen beruflichen Weiterentwicklung an.
- Das Feedback wird vom Empfangenden als relevant und sinnvoll betrachtet.
- Beim Vorbereiten und Darbieten des Feedbacks wurden der Erfahrungsschatz und Entwicklungsstand des Empfangenden berücksichtigt.
- Feedback ist wirksamer, wenn es auf das Selbstwertgefühl des Empfängers eingeht.
- Feedback ist effektiver, wenn es die Unsicherheit des Empfängers verringert.

Inhalt des Feedbacks (Was?)

- Das Feedback bezieht sich auf Ziele, die von der empfangenden Person definiert wurden, oder in einer positiven Leistung resultieren.
- Das Feedback enthält genaue Daten und eindeutige Beweise, konkrete Informationen und Beispiele für ein erwünschtes Ergebnis.
- Das Feedback ist zielgerichtet, beschreibend, nicht bewertend und fokussiert auf das Verhalten und nicht auf die Person.
- Das Feedback enthält einen moderaten Anteil an positivem Feedback und einen ausgewählten und begrenzten Anteil an negativem Feedback.
- Das negative Feedback ist zwischen das positive Feedback eingebettet.
- Das Feedback erzeugt ein moderates Maß an kognitiver Dissonanz.
- Der Feedbackprozess ermöglicht eine Reaktion und Interaktion.

Art des Feedbacks (Wie?)

- Das Feedback wird auf verschiedene Formen vermittelt.
- Die empfangende Person hat die Möglichkeit, die Art und Weise des Feedbacks, wie es übermittelt wird, selbst zu wählen.

Dauer des Feedbacks (Wann?)

- Das Feedback wird so schnell wie möglich nach der Leistung gegeben.
 - Das Feedback wird als Prozess betrachtet und nicht als einmalige schnelle Lösung.
 - Das Feedback wird häufig, aber nicht mehr als notwendig gegeben.
-

Empirische Befunde zur Wirksamkeit von Feedback speziell im Vorbereitungsdienst oder ähnlichen Vorbereitungskursen im internationalen Sprachraum sind wenig vorhanden (vgl. Bitan, 2014). So gibt es beispielsweise eine amerikanische qualitative Studie von Wilcoxon und Lemke (2021), in der die Erwartungen von Lehrkräften berichtet werden, welche ein vierjähriges Studium abgeschlossen haben und sich in dem ersten Jahr ihrer Lehrtätigkeit befinden. Die Befunde unterscheiden sich nicht von den bisher geschilderten Implikationen in Kapitel 2.3.2, sondern bestätigen diese. Die teilnehmenden Lehrkräfte wünschen sich ausführliches, hochwertiges, konstruktives und regelmäßiges Feedback zu Zeitpunkten, zu denen es noch nicht zu spät ist, um das erhaltene Feedback in der Unterrichtspraxis umzusetzen. Zudem wünschen sich die Lehrkräfte Feedback, das ihnen mehr Hinweise und Einsichten zur Verbesserung ihres Unterrichts gibt. Ein Feedback, das nur die positiven Aspekte hervorhebt, würde keinen Raum für Verbesserung einräumen und sei daher nicht nützlich. Die Teilnehmer weisen außerdem auf die Bedeutung des schriftlichen Feedbacks hin, da es ihnen die Möglichkeit gäbe, zunächst zu reflektieren und nachträglich zu reagieren (Wilcoxon & Lemke, 2021).

2.4 Zusammenfassung des bisherigen Forschungsstandes und Desiderate

Der theoretische Hintergrund der vorliegenden Arbeit beleuchtet zentrale Einflussfaktoren auf die Professionalisierung angehender Chemielehrkräfte in Bezug auf ihre repräsentationsbezogene Kompetenz. Es wird deutlich, dass die Qualität des Chemieunterrichts und der eingesetzten Lehr-Lern-Materialien maßgeblich davon abhängt, inwieweit Lehrkräfte in der Lage sind, fachliche Inhalte in geeigneten externen Repräsentationen aufzubereiten und lernförderlich zu vermitteln. Außerdem ist das Unterrichtsfach Chemie in besonderem Maße gekennzeichnet durch abstrakte Konzepte, symbolische Darstellungen und Modelle, die für Schüler:innen zumeist schwer zugänglich sind. Zudem werden im Chemieunterricht in der Regel mehrere externe Repräsentationen in Kombination eingesetzt, um chemische Phänomene verständlich zu machen, was Schüler:innen überfordern kann.

Für angehende Lehrkräfte bedeutet dies, dass sie sowohl über ein fundiertes repräsentationsbezogenes Fachwissen als auch ein repräsentationsbezogenes fachdidaktisches Wissen verfügen müssen. Während sich das Fachwissen auf das sachgerechte Interpretieren, Konstruieren und fachliche Bewerten von Repräsentationen bezieht, fokussiert das fachdidaktische Wissen den lernwirksamen Einsatz von Repräsentationen in konkreten Unterrichtssituationen und die Berücksichtigung von Lernschwierigkeiten der Schüler:innen auf fachdidaktischer und kognitionspsychologischer Ebene.

Neben den kognitiven Wissensbeständen sind auch affektive Handlungskompetenzen wie die repräsentationsbezogenen Überzeugungen zu berücksichtigen. Repräsentationsbezogene Überzeugungen beinhalten die Einschätzung der Nützlichkeit und Verständlichkeit naturwissenschaftlicher externer Repräsentationen sowie die persönliche Haltung der Lehrkräfte gegenüber ihrer Verwendung im Chemieunterricht. Da repräsentationsbezogene Überzeugungen das Handeln von Lehrkräften steuern, das die Auswahl und Gestaltung von Repräsentationen betrifft, haben sie einen erheblichen Einfluss auf das unterrichtspraktische Handeln. Eine Veränderung oder Weiterentwicklung der Überzeugungen wäre daher von Bedeutung für eine langfristige Professionalisierung.

Ein zentrales Problem besteht jedoch darin, dass die genannten Wissensbereiche und Überzeugungen in der Lehrkräftebildung kaum systematisch gefördert werden und den angehenden Lehrkräften Lerngelegenheiten fehlen, in denen sie das Gestalten und Reflektieren multipler externer Repräsentationen und Unterrichtsmaterialien einüben können. Noch weniger ist über die gezielte Förderung solcher Lernprozesse während des Vorbereitungsdienstes durch Feedback bekannt, obwohl Feedback in der gesamten Lehr-Lern-Forschung als einer der wirksamsten Faktoren für Lernerfolg und Entwicklung gilt. Es bleibt die Frage offen, welchen Einfluss Feedback auf das repräsentationsbezogene Fachwissen und fachdidaktische Wissen sowie die repräsentationsbezogenen Überzeugungen hat.

Darüber hinaus lässt sich aus dem Forschungsstand ableiten, dass systematisches internes und externes Feedback in der Aus- und Weiterbildung von Lehrkräften besonders förderlich ist. Bisher fehlen jedoch Studien, die die Effekte der beiden Feedbackquellen sowie den Einfluss auf das repräsentationsbezogene Fachwissen, fachdidaktische Wissen und die Überzeugungen untersuchen. Ebenso ist unklar, ob es differentielle Effekte zwischen dem internen und externen Feedback gibt.

Vor diesem Hintergrund ergibt sich die leitende Fragestellung, die sowohl an das Desiderat der Feedbackforschung als auch den fachspezifischen Bedarf der Chemiedidaktik anknüpft. Die hieraus resultierende und vorliegende Studie trägt somit zur Forschung im Bereich der Professionalisierung von Lehrkräften in der zweiten Phase der Lehramtsausbildung bei, wie sie etwa von Anderson-Park und Abs (2020), Behrens et al. (2023) sowie Hascher (2014) gefordert wird.

3 ZIEL, FORSCHUNGSFRAGEN UND HYPOTHESEN

Im Anschluss an die theoretischen Überlegungen soll mit dem vorliegenden Forschungsvorhaben untersucht werden, wie eine feedbackgestützte Intervention das Professionswissen sowie die Überzeugungen angehender Lehrkräfte des Unterrichtsfaches Chemie zu von ihnen entworfenen multiplen externen Repräsentationen beeinflusst. Der Fokus liegt dabei einerseits auf der Wirkung der internen und externen Feedbackquellen und andererseits auf der Entwicklung des repräsentationsbezogenen Fachwissens, fachdidaktischen Wissens und der Überzeugungen. Zudem wird geprüft, ob und inwieweit der Effekt von Feedback durch das Vorwissen und durch individuelle Voraussetzungen und Bedingungen der angehenden Lehrkräfte beeinflusst wird.

Um den in Kapitel 2.4 beschriebenen Forschungsbedarf und das Untersuchungsziel zu adressieren, werden im Rahmen der vorliegenden Dissertation folgende Forschungsfragen hinsichtlich der Wirkung von Feedback aufgestellt:

- F_{1A}: Welchen Effekt haben internes und/oder externes Feedback auf das repräsentationsbezogene Fachwissen und fachdidaktische Wissen von angehenden Lehrkräften des Faches Chemie?**
- F_{1B}: Welchen Effekt haben internes und/oder externes Feedback auf die Überzeugungen angehender Lehrkräfte des Faches Chemie in Bezug auf die Relevanz von externen Repräsentationen für das Lernen von Chemie?**
- F_{1C}: Welchen Effekt haben internes und/oder externes Feedback auf die Qualität der von angehenden Lehrkräften designten multiplen externen Repräsentationen?**
- F_{1D}: Welchen Effekt haben internes und/oder externes Feedback auf die Selbstbewertungsfähigkeit von angehenden Lehrkräften des Faches Chemie in Bezug auf selbstdesignte multiple externe Repräsentationen?**

Um die spezifischen Wirkungen des jeweils internen und externen Feedbacks fundiert zu untersuchen, ist Vergleich zwischen differenzierten feedbackgestützten Interventionen erforderlich. Nur durch eine Differenzierung können Daten

und Erkenntnisse darüber gewonnen werden, wie internes und externes Feedback jeweils einzeln sowie in Kombination wirken. Auf Grundlage der bestehenden Forschungslage zu Feedback (Kapitel 2.3) lässt sich dennoch grundsätzlich annehmen, dass der Einsatz von Feedback im Hinblick auf die genannten abhängigen Variablen einen positiven und förderlichen Effekt hat. Es ist davon auszugehen, dass alle teilnehmenden angehenden Lehrkräfte Fortschritte erzielen, da alle differenziert geplanten Interventionen das Arbeiten mit Feedback vorsehen, unabhängig von der jeweils eingesetzten Feedbackquelle. Auch Taskin et al. (2015a) verdeutlichen, dass sich das repräsentationsbezogene Fachwissen und fachdidaktische Wissen im Laufe des Studiums allein durch die Auseinandersetzung mit Repräsentationen erhöht, selbst wenn externe Repräsentationen kein expliziter Lerngegenstand sind. Es ist daher sogar anzunehmen, dass auch eine Kontrollgruppe, die ohne systematisches internes und/oder externes Feedback interveniert wird, einen Lernzuwachs aufweist. Es wird jedoch erwartet, dass eine Intervention mit internem und/oder externen Feedback im Vergleich einen signifikanten Vorteil zeigt.

Bezogen auf die Differenzierung zwischen den unterschiedlichen Feedbackbedingungen, ob internes und externes Feedback jeweils einzeln oder in Kombination effektiver ist, wird nach dem ITFL-Modell von Narciss (2018) angenommen, dass die ohnehin immer stattfindende interne Feedbackschleife durch extern dargebotene Ist- und Soll-Werte unterstützt wird, wodurch der gesamte Feedbackprozess und damit auch der Lernprozess verstärkt wird. Demnach wird angenommen, dass die Kombination von internem und externem Feedback einen größeren Effekt hat als der alleinige Einsatz von internem Feedback. Zusammengefasst können folgende Hypothesen hergeleitet werden:

H_{1A.1}: Eine Intervention mit internem und/oder externem Feedback hat einen größeren Effekt auf beide Wissensarten als eine Intervention ohne Feedback.

H_{1A.2}: Eine Intervention mit internem und externem Feedback hat einen größeren Effekt auf beide Wissensarten als eine Intervention nur mit internem Feedback.

Überzeugungen von Lehrkräften bezüglich der Relevanz externer Repräsentationen für das Lernen im Chemieunterricht gelten zwar als vergleichsweise zeitlich stabile Strukturen, sind jedoch durch gezielte Impulse veränderbar (Kunter et al., 2013). Feedback kann diese Impulse geben, vor allem wenn es wie in der vorliegenden Studie spezifisch, konstruktiv und auf individuelle Lernfortschritte ausgerichtet ist. Dadurch können auch Reflexionsprozesse angeregt werden, die die affektiven Merkmale der Feedbackerhaltenden ansprechen, ausdifferenzieren und fördern (Hattie & Timperley, 2007). Darüber hinaus zeigen empirische Studien, dass insbesondere formatives Feedback das Metadenken und die Selbstreflexion fördert, was wiederum zu einer Verstärkung der Überzeugungen beitragen kann (Narciss, 2018; Wisniewski et al., 2019). Daher ist davon auszugehen, dass feedbackgestützte Interventionen neben dem kognitiven Wissenszuwachs auch einen förderlichen Effekt auf die Überzeugungen der angehenden Lehrkräfte entfalten.

H_{1B.1}: Eine Intervention mit internem und/oder externem Feedback hat einen größeren Effekt auf die Überzeugungen als eine Intervention ohne Feedback.

H_{1B.2}: Eine Intervention mit internem und externem Feedback hat einen größeren Effekt auf die Überzeugungen als eine Intervention nur mit internem Feedback.

Bei der Qualität der designten multiplen Repräsentationen wird angenommen, dass ein höheres Maß an repräsentationsbezogenem Fachwissen und fachdidaktischem Wissen sowie stärkeren Überzeugungen der angehenden Lehrkräfte auch dazu beitragen, Details in externen Repräsentationen genauer zu erkennen und fundierter zu bewerten. Gleichzeitig wird vermutlich auch die Fähigkeit gefördert, konkrete Optimierungen und Alternativen der Gestaltung zu entwickeln und umzusetzen. Darüber hinaus erhalten die angehenden Lehrkräfte im Rahmen einer feedbackgestützten Intervention konkretes und individuelles Feedback zu ihren designten multiplen externen Repräsentationen. Dadurch erwerben sie zusätzlich grundlegende fachdidaktische Prinzipien für die Gestaltung multipler externer Repräsentationen, die sie langfristig anwenden und umsetzen können.

H_{1C.1}: Eine Intervention mit internem und/oder externem Feedback hat einen größeren Effekt auf die Qualität der designten multiplen externen Repräsentationen als die Intervention ohne Feedback.

H_{1C.2}: Eine Intervention mit internem und externem Feedback zeigt einen größeren Effekt auf die Qualität als die Intervention nur mit internem Feedback.

Die Selbstbewertungsfähigkeit beschreibt die Kompetenz der angehenden Lehrkräfte, ihre selbst designten multiplen externen Repräsentationen realistisch und zutreffend einzuschätzen, sodass sie sich eigenständig internes Feedback geben können, das nach dem ITFL-Modell möglichst genau dem tatsächlichen Ist-Zustand entspricht. Diese Fähigkeit kann insbesondere dann gefördert werden, wenn die angehenden Lehrkräfte zunächst ihr eigenes internes Feedback generieren und dieses anschließend mit einem direkt vergleichbaren externen Feedback, welches als Maßstab für den Ist-Wert dient, abgleichen. Durch die wiederholte Durchführung dieses Prozesses können die angehenden Lehrkräfte ein präziseres Einschätzungsvermögen entwickeln, als wenn ausschließlich internes Feedback ohne ergänzendes externes Feedback genutzt würde.

H_{1D}: Eine Intervention mit internem und externem Feedback zeigt einen größeren Effekt auf die Selbstbewertungsfähigkeit als die Intervention mit nur internem Feedback.

Um auch den Einfluss der individuellen Voraussetzungen der angehenden Lehrkräfte zu untersuchen, wurde folgende zweite Forschungsfrage aufgestellt:

F₂: Welchen Effekt haben allgemeine Personenmerkmale der angehenden Lehrkräfte des Faches Chemie auf das repräsentationsbezogene Fachwissen und fachdidaktische Wissen?

Zur Herleitung von Hypothesen können Befunde aus den Arbeiten von Dollny (2011), Mutke (2017) und Taskin et al. (2015a) herangezogen werden, die auf den signifikanten Einfluss der Schulform hinweisen. Dies betrifft nicht nur das allgemeine Fachwissen und fachdidaktische Wissen langjähriger Lehrkräfte

(Dollny, 2011), sondern auch das der angehenden Lehrkräfte im Vorbereitungsdienst (Mutke, 2017). Die Unterschiede im Fachwissen von Lehrkräften unterschiedlicher Schulformen bzw. -stufen spiegeln sich auch im repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen von Lehramtsstudierenden wieder (Taskin et al., 2015a, 2015b). Die berichteten Unterschiede beziehen sich insbesondere auf Unterschiede im Professionswissen von Lehrkräften gymnasialer und nicht-gymnasialer Schulformen. Chemielehrkräfte gymnasialer Schulformen performen im fachbezogenen Professionswissen besser als Chemielehrkräfte nicht-gymnasialer Schulformen.

Zusätzlich liegen Hinweise vor, dass ein höheres allgemeines Fachwissen auch mit einem Anstieg des fachdidaktischen Wissens einhergeht und umgekehrt ($r = .32$). Beide Kompetenzbereiche wirken dabei als gegenseitige Prädiktoren füreinander (Kapitel 2.2.2). Gleiche Effekte sind auch beim repräsentationsbezogenen Fachwissen und beim repräsentationsbezogenen fachdidaktischen Wissen erwartbar.

Ein weiteres zentrales Thema, welches oft in der Literatur – auch fächerübergreifend – erforscht wird (z. B. Gehrman, 2023; Porsch et al., 2023), ist die Frage, ob Seiteneinsteiger:innen oder angehende Lehrkräfte, die über einen Abschluss eines gleichwertigen Studiums ohne Lehramtsoption den Zugang zum Vorbereitungsdienst erhalten, anders performen als grundständige Lehramtsanwärter:innen. Da frühere Befunde diesbezüglich keine signifikanten Unterschiede festgestellt haben (z. B. Dollny, 2011; Mutke, 2017), wird erwartet, dass keine Unterschiede bestehen. Vor dem Hintergrund können die folgenden Hypothesen formuliert werden:

H_{2.1}: Das fachbezogene Professionswissen von angehenden Lehrkräften, die den Vorbereitungsdienst für die Sekundarstufe I und II absolvieren, ist höher als von angehenden Lehrkräften, die den Vorbereitungsdienst nur für die Sekundarstufe I absolvieren.

H_{2.2}: Angehende Lehrkräfte mit einem hohen repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen lernen mehr dazu als angehende Lehrkräfte mit einem niedrigen repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen.

H_{2,3}: Das fachbezogene Professionswissen von angehenden Lehrkräften, die ein Studium mit Lehramtsoption abgeschlossen haben, ist genauso hoch, wie das fachbezogene Professionswissen von angehenden Lehrkräften, die ein Studium ohne Lehramtsoption abgeschlossen haben.

Um die genannten Forschungsfragen und Hypothesen zu untersuchen, wurden in der vorliegenden Arbeit, wie im folgenden Methodenteil (Kapitel 4) dargestellt, feedbackgestützte Interventionen entwickelt und erprobt. Gegenstand der Untersuchung ist ein experimenteller Ansatz zur Förderung des Wissens und der Überzeugungen angehender Lehrkräfte. Ziel des Forschungsvorhabens ist es, die Verknüpfung von Theorie und Praxis zu stärken. Ein zentrales Element dieses Ansatzes ist das Feedback und die Variation der Quelle des Feedbacks. Die zu erwartenden Ergebnisse dieser Untersuchung (Kapitel 5 und 6) sollen dazu beitragen, die Lehrkräfteausbildung in der zweiten Phase der Lehramtsausbildung zu verbessern.

4 METHODISCHES VORGEHEN

Zur Untersuchung der aufgestellten Forschungsfragen und hergeleiteten Hypothesen wurde eine experimentelle Interventionsstudie im Prä-Post-Kontrollgruppendesign entwickelt und mit angehenden Lehrkräften im Vorbereitungsdienst im Bundesland Nordrhein-Westfalen erprobt. Das Studiendesign, die Inhalte der Intervention, die eingesetzten Erhebungsinstrumente sowie relevanten Entwicklungsschritte werden im Folgenden erläutert.

4.1 Studiendesign

4.1.1 Untersuchungsvariablen

Als unabhängige Variable werden die Feedbackquellen betrachtet. Dabei handelt es sich zum einen um intern generiertes Feedback, bei dem sich Proband:innen selbst ein Feedback geben und zum anderen um extern generiertes Feedback, bei dem die Proband:innen ein Feedback von Seiten der Studienleitung erhalten. Idealerweise ergibt die systematische Variation und Kombination der beiden Feedbackquellen ein 2x2-Interventionsdesign mit vier Gruppen: Internes und externes Feedback, internes Feedback, externes Feedback, kein Feedback. Prognosen zur Grundgesamtheit aller angehenden Lehrkräfte im Vorbereitungsdienst des Faches Chemie im Bundesland Nordrhein-Westfalen während der Planung ergaben jedoch, dass die für valide und reliable Ergebnisse erforderliche Mindeststichprobengröße von 30 Proband:innen je Untersuchungsgruppe nicht erreicht werden kann (Kapitel 4.1.2). Infolge wurde das Studiendesign angepasst und um eine Untersuchungsgruppe reduziert. Bei der theoriegeleiteten Abwägung, welche der ursprünglich vier Untersuchungsgruppen am ehesten ausgespart werden könnte, wurde die Untersuchungsgruppe „externes Feedback“ ausgewählt und ausgeschlossen. Da beim Geben externen Feedbacks unbewusst und unvermeidbar auch interne Feedbackprozesse und Reflexionen generiert werden, ist ohnehin schwer begründbar, dass diese Gruppe ausschließlich mit externem Feedback arbeitet (z. B. nach Narciss, 2018; Kapitel 2.3.1). Eine Treatmentkontrolle und ein trennscharfer Vergleich zwischen den Feedbackbedingungen wären daher als eingeschränkt zu bewerten. Beide Aspekte sind in den anderen drei Untersuchungsgruppen stärker gewährleistet, sodass folgende umgesetzt wurden: Interventionsgruppe 1 (IG1, internes

und externes Feedback), Interventionsgruppe 2 (IG2, internes Feedback), Kontrollgruppe (KG, kein Feedback). Details über die zugehörigen Interventionen in den drei Untersuchungsgruppen sind in Kapitel 4.1.3 dargestellt.

Als abhängige Variablen fokussiert die Studie das repräsentationsbezogene Fachwissen (Rep. FW), das repräsentationsbezogene fachdidaktische Wissen (Rep. FDW) sowie die repräsentationsbezogenen Überzeugungen (Rep. Ü). Die Auswahl und Adaption der entsprechenden Testinstrumente werden in den Kapiteln 4.2.3 und 4.2.4 ausgeführt. Des Weiteren wurden während der Intervention der wahrgenommene Cognitive Load und die wahrgenommene Anstrengungsbereitschaft erfasst. Diese lassen sich einerseits als weitere abhängige Variablen interpretieren und bieten zusätzliche Einblicke in die Arbeitsprozesse und Entwicklungsverläufe der Proband:innen (Kapitel 4.2.5).

Als Kontrollvariablen wurden das allgemeine fachliche und fachdidaktische Wissen erhoben. Die Kontrolle des allgemeinen, nicht repräsentationsbezogenen Wissens erlaubt eine Aussage über mögliche Zusammenhänge zwischen allgemeinem und repräsentationsbezogenem Wissen und dient der diskriminanten Validierung. Ebenso kann dadurch untersucht werden, ob die Kompetenzen einen prädiktiven Einfluss auf die Effektivität des Feedbacks haben.

Als weitere Kontrollvariablen wurden neben personenbezogenen Variablen, Angaben zum Ausbildungshintergrund und zur aktuellen beruflichen Situation auch individuelle Bedingungen, wie beispielsweise zeitliche Ressourcen und Belastungen, sowie die Beurteilung von Ausbildungstätigkeiten gemessen (Kapitel 4.1.2). Des Weiteren wurden das professionelle Selbstverständnis, darunter Aspekte des Wissenschaftsverständnisses und die berufliche Selbstwirksamkeit, erfasst (Abs et al., 2005). Um eine mögliche Abhängigkeit der Wirksamkeit der Intervention von der Fähigkeit der Proband:innen, eigene Lernmaterialien zu konzipieren und zu kontrollieren, wurde die Erfahrung der Proband:innen in der Erstellung von Lernmaterialien durch Selbsteinschätzung als Kontrollvariable erhoben.

Tabelle 4: Untersuchungsvariablen der Interventionsstudie

Unabhängige Variablen	Abhängige Variablen	Kontrollvariablen
<i>Intervention:</i>	<i>Prä- und Posttest:</i>	<i>Prätest:</i>
- IG2: Internes Feedback per Bewerbungsbogen	- Repräsentationsbezogenes Fachwissen	- Demografische Daten (Alter, Muttersprache(n), Geschlecht, Note des Abschlusses im Abitur, Lehramtsform, Note des universitären Abschlusses, Schulform, Dienstgrad, absolvierte Zeit im Vorbereitungsdienst, Dauer einer Vertretungstätigkeit vor dem Vorbereitungsdienst)
- Internes und externes Feedback per Bewertungsbogen	- Repräsentationsbezogenes fachdidaktisches Wissen	- Persönliche Ressourcen in Bezug auf den Vorbereitungsdienst (Belastungen und Bewältigungsstrategien, Professionelles Selbstverständnis, Erfahrung in der Gestaltung von Lehr-Lern-Materialien)
- Dummy-Feedback per kurzem Text	- Repräsentationsbezogene Überzeugungen	
	<i>Interventionsbegleitend:</i>	
	- Qualität der Lehr-Lern-Materialien	
	- Qualität des internen Feedbacks	
	- Cognitive Load	
	- Anstrengungsbereitschaft	
		<i>Prä- und Posttest:</i>
		- Allgemeines Fachwissen
		- Allgemeines fachdidaktisches Wissen

Anmerkung. Lehramtsform: HRSGe, GymGe; Schulform: Hauptschule, Realschule, Sekundarschule, Gesamtschule, Gymnasium; Dienstgrad: Seiteneinsteiger:in o. ä., Lehramtsanwärter:in.

4.1.2 Akquise und Gruppenzuordnung der Stichprobe

Im Rahmen der Intervention wurde die Teilnahme von insgesamt $N = 120$ angehenden Lehrkräften im Vorbereitungsdienst der beiden in Nordrhein-Westfalen für den Sekundarbereich wählbaren Lehramtsformen (Haupt-/Real-/Sekundar-/Gesamtschule, Gymnasium/Gesamtschule) angestrebt. Der Vergleich zwischen den gymnasialen und nicht-gymnasialen Schulformen zielt darauf ab, Aufschluss darüber zu gewinnen, ob unterschiedliche Methodenkompetenzen oder auch Studiengänge einen Einfluss auf das repräsentationsbezogene Wissen haben. Die Rekrutierung der Proband:innen strebte die Einbeziehung aller angehenden Lehrkräfte im Vorbereitungsdienst an, die das Fach Chemie als eines ihrer Unterrichtsfächer haben. Mit der Vollerhebung in Nordrhein-Westfalen solle eine möglichst hohe Repräsentativität der Studienergebnisse gewährleistet werden (nach Döring & Bortz, 2023). Die dazu erforderliche Genehmigung wurde zuvor vom Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen eingeholt. Für die Teilnahme an der Intervention wurde den angehenden Lehrkräften eine Aufwandsentschädigung in Höhe von 100 Euro entrichtet. Um Selektionseffekte bei der Zusammenstellung der Untersuchungsgruppen zu vermeiden, erfolgte die Zuweisung der angehenden Lehrkräfte zu den drei Untersuchungsgruppen in drei nacheinander folgenden Schritten (nach Wayne et al., 2008): Im ersten Schritt wurden die Proband:innen nach ihrer studierten Lehramtsform (HRSGe, GymGe) aufgeteilt. Innerhalb dieser Einteilung wurde im zweiten Schritt unterschieden, ob die Proband:innen einen klassischen Lehramtsabschluss (Lehramt) oder einen anderen Hochschulabschluss (Seiteneinstieg) aufweisen. Daraus ergaben sich insgesamt vier Untergruppierungen, die im letzten Schritt randomisiert gleichmäßig an eine der drei Untersuchungsgruppen zugewiesen wurden. Die hierzu erforderlichen Variablen wurden im Anmeldebogen erfragt.

Dadurch wurde bereits im Vorfeld der Intervention darauf geachtet, die Verteilung so vorzunehmen, dass einerseits Clusterbildungen vermieden und andererseits personenbezogene Störvariablen kontrolliert werden. Weitere potenzielle Störvariablen, wie beispielsweise das Fachwissen oder die Abschlussnoten, wurden im Prätest erfasst und bei der Stichprobencharakterisierung sowie bei den Ergebnisanalysen berücksichtigt (Kapitel 5.1).

4.1.3 Struktur und Inhalt der drei Online-Interventionssitzungen

Die Intervention besteht aus drei aufeinander aufbauenden Online-Sitzungen, die in den regulären Vorbereitungsdienst der angehenden Lehrkräfte integriert wurden (Tabelle 5). Vor und nach der Intervention fanden Erhebungen statt, um die Kontrollvariablen und abhängigen Variablen zu erfassen. Die zugehörigen Testinstrumente sind in Kapitel 4.2 dargestellt.

Tabelle 5: Interventionsdesign

Untersuchungs- gruppe	1. Prätest (105 min)	2. Intervention (3 x 90 min)	3. Posttest (80 min)
Interventions- gruppe 1 (IG1, $n = 29$)	<i>Kontrollvariablen</i> - Demographische Daten - Persönliche Ressourcen	Internes und externes Feedback	<i>Kontrollvariablen</i>
Interventions- gruppe 2 (IG2, $n = 29$)	- Allg. Fachwissen und fachdidaktisches Wissen	Internes Feedback	- Allg. Fachwissen und fachdidaktisches Wissen
Kontroll- gruppe (KG, $n = 32$)	<i>Abhängige Variablen</i> - Rep. Fachwissen und fachdidaktisches Wissen - Rep. Überzeugungen	Dummy Feedback	<i>Abhängige Variablen</i> - Rep. Fachwissen und fachdidaktisches Wissen - Rep. Überzeugungen

Die Struktur der Intervention war in allen drei Untersuchungsgruppen gleich, unterschied sich jedoch inhaltlich in den Feedbackphasen der Intervention. In Anlehnung an die Implikationen von Reusser und Pauli (2014) über die Aus- und Weiterbildung von Lehrkräften wurde bei der Intervention eine Kombination aus exemplarischen Inputphasen sowie unterrichtsbezogenen Übungs- und Feedbackphasen angewandt. Daher bestand jede Interventionssitzung aus vier Phasen (Tabelle 6): 1. *Einleitung* bzw. Aufgabenstellung, 2. *Übung* bzw. Bear-

beitung der Aufgabe, 3. Generierung des *Feedbacks*, 4. *Reflexion* bzw. Verarbeitung des Feedbacks. Dabei unterschieden sich die drei Untersuchungsgruppen in der Durchführung der Feedbackphasen (3., s. u.). Die Einleitungs-, Übungs- und Reflexionsphasen verliefen für jede Untersuchungsgruppe in jeder Sitzung methodisch und strukturell identisch. Es wurde jedoch in jeder Sitzung ein anderer chemischer Lerninhalt behandelt (Kapitel 4.1.4).

Die Aufgabe der angehenden Lehrkräfte bestand darin, ein Lehr-Lern-Material zu einem vorgegebenen Lernziel zu designen und anschließend im Rahmen einer Feedbackphase darüber zu reflektieren, mit dem Ziel, in der nächsten Sitzung ein besseres Design für ein weiteres Lehr-Lern-Material umsetzen zu können. Um die Authentizität der Intervention zu erhöhen, wurde darauf verzichtet, die angehenden Lehrkräfte, lediglich eine multiple externe Repräsentation isoliert als solche entwerfen zu lassen. Stattdessen wurden die multiplen externen Repräsentationen in konkrete unterrichtliche Situation eingebettet. Die angehenden Lehrkräfte erhielten hierfür als Ausgangspunkt eine festgelegte Klassenstufe, ein spezifisches Lernziel sowie weitere Angaben zur Lerngruppe. Auf dieser Grundlage sollten sie ein Lehr-Lern-Material entwickeln, das nicht nur multiple externe Repräsentationen umfasst, sondern auch einen schüler:innengerechten Eingangskontext sowie passende Aufgabenstellungen enthält.

Um die Realisierbarkeit der Studie während der Corona-Pandemie zu ermöglichen, wurden sowohl die Interventionssitzungen als auch der Prä- und Posttest in einen Moodle-Kurs implementiert (Anhang A.1). Alle Interventionsmaterialien waren darin einseh- und bearbeitbar. Auch die Durchführungsobjektivität der Studie konnte so erhöht werden. Ferner konnte auf eine Schulung bzw. ein Monitoring der Studienleitung verzichtet werden. Somit waren die Proband:innen in der Lage ihre Sitzungen individuell und flexibel zu terminieren. Ihnen wurde ermöglicht, die drei Sitzungen in selbstgewählten Abständen von jeweils fünf bis sieben Tagen zu absolvieren.

In Absprache mit den Fachseminarleitungen wurde die Intervention als eigenständig bearbeitbares und interaktives Modul in den regulären Vorbereitungsdienst integriert. Die Intervention wurde seitens der Fachleitungen für die angehenden Lehrkräfte entweder als verpflichtend und obligatorisch oder als freiwilliges Zusatzangebot angeboten. Da die Intervention als Online-Angebot während des Beginns der Pandemie und der damit einhergehenden Schul- und

Seminarschließungen zur Verfügung gestellt wurde, diente sie in erster Linie als Ersatz für ausfallende Fachseminarsitzungen. Infolgedessen konnte innerhalb eines Erhebungszeitraums von etwa zwei Monaten eine verhältnismäßig hohe Stichprobe erreicht werden.

Die einzelnen Interventionssitzungen wurden im Moodle-Kurs so gestaltet, dass die o. g. Phasen nacheinander bearbeitbar sind (Tabelle 6). Zur Einleitung wurde den angehenden Lehrkräften eine interaktive Präsentation mittels der Software H5P dargeboten. Darin wurden die Proband:innen in die Thematik eingeführt, warum es aus (fach-)didaktischer und kognitionspsychologischer Sicht von Bedeutung ist, auf Details in (multiplen) Repräsentationen zu achten und diese bewusst an die Lernenden anzupassen. Anschließend wurden die Proband:innen an ihre Übungsaufgabe herangeführt. Ihnen wurden ein festes Lernziel für eine Lerngelegenheit, dazu erforderliche Text- und Bildmaterialien und konkrete Rahmenbedingungen einer fiktiven Lerngruppe vorgegeben (Kapitel 4.1.4, Anhang A.2). Aus diesen Materialien sollten sie in der Übungsphase in einer Zeitdauer von 45 Minuten ein fertiges Lehr-Lern-Material entwerfen.

Nach Anfertigung des Lehr-Lern-Materials sicherten die Proband:innen ihr designtes Lehr-Lern-Material und luden es auf der Moodle-Plattform hoch, um im Anschluss in die Feedbackphase durchzuführen. Zunächst wurde je nach Untersuchungsgruppe ein Feedback über den Ist-Stand des Lehr-Lern-Materials generiert. Auf der fünfstufigen Likert-Skala (Kapitel 4.2.5) sahen die Testpersonen gleichzeitig, auf welchem Niveau sie sich innerhalb der Skala befinden. Im Anschluss wurden die Feedbackinformationen von den Proband:innen reflektiert. Die Proband:innen sollten die Informationen nutzen, um in der darauffolgenden Sitzung bzw. Übung ein Lehr-Lern-Material anzufertigen, das den Anforderungen noch besser entspricht. Dieser Zyklus von Übung, Feedbackgenerierung und Reflexion wurde dreimal durchgeführt.

In Anhang A.1 sind Ausschnitte aus dem Moodle-Kurs abgebildet und es ist erkennbar, wie die Testungen und die einzelnen Interventionssitzungen inkl. Phasen realisiert wurden. Die angehenden Lehrkräfte wurden anhand möglichst klar formulierter Arbeits- und Handlungsanweisungen durch die Intervention geführt. Durch das Anklicken der Elemente konnten entsprechende Informationen geöffnet, Dateien hoch- und heruntergeladen sowie Eingabemasken bearbeitet werden.

Tabelle 6: Ablauf und Struktur einer Interventionssitzung am Beispiel der ersten Sitzung

<i>Phase</i>	<i>Tätigkeiten der angehenden Lehrkräfte</i>
Einleitung 1	- Einführung in das Ziel der Intervention
	- Aufgabenstellung und Ausgabe der Übungsaufgabe mit vorgegebenem Lernziel 1 und Rohmaterial 1
Übung 1	- Gestaltung eines Lehr-Lern-Materials 1
Feedback-generierung 1	- <i>IG1</i> : Erhalt des Bewertungsbogens und Generierung des internen Feedbacks per Bewertungsbogen für das gestaltete Lehr-Lern-Material 1
	- <i>IG2</i> : Erhalt des Bewertungsbogens und Generierung des internen Feedbacks per Bewertungsbogen für das gestaltete Lehr-Lern-Material 1
	- <i>KG</i> : Erhalt von themenidentischen Schulbuchauszügen und Vergleich mit gestaltetem Lehr-Lern-Material 1
	- <i>IG1</i> : Erhalt des externen Feedbacks per Bewertungsbogen, Vergleich des internen Feedbacks mit externem Feedback
Reflexion 1	- <i>IG2</i> : Erhalt des Dummy-Feedbacks, Vergleich des internen Feedbacks mit dem Dummy-Feedback
	- <i>KG</i> : Erhalt des Dummy-Feedbacks

Bezogen auf die Übungsphase durchliefen alle drei Untersuchungsgruppen die zentralen Schritte zur Erstellung eines Lehr-Lern-Materials: Ausgehend von einer konkreten Lerngruppe sollte ein Lernziel aus dem Kernlehrplan oder Schulcurriculum ausgewählt werden. Dies wurde in der Intervention vorgegeben. Danach sollen geeignete Darstellungsformen für den Einsatz im Lehr-Lern-Material recherchiert werden. Eine kleine Sammlung geeigneter Texte und Abbildungen wurde in der Intervention hierfür vorgegeben. Aus diesen einzelnen externen Repräsentationen wurde ein ganzheitliches und einsatzbereites Lehr-Lern-Material, bzw. eine multiple externe Repräsentation, gestaltet. Dabei wurden die angehenden Lehrkräfte angeleitet, die Abbildungen und Texte nach ihrem eigenen Einschätzungsvermögen auszuwählen, zu überarbeiten und anzupassen, geeignete Aufgabenstellungen zu ergänzen und das Arbeitsblatt mit

weiterer Detailarbeit zu finalisieren. Da der gesamte Prozess dreimal durchlaufen wurde, beschränkte sich die Intervention nicht nur auf das Lernen, sondern führte auch zur Routinierung des Designprozesses.

Im Folgenden werden die methodischen Entscheidungen bezüglich der Feedbackgenerierung und Reflexion dargestellt, die sich zwischen den Untersuchungsgruppen unterschieden. Die beiden Interventionsgruppen arbeiteten mit systematischem internem und/oder externem Feedback, während die Kontrollgruppe mit keinem systematischen internem oder externem Feedback arbeitete. Zur systematischen Durchführung des Feedbacks in den Interventionsgruppen nutzten die angehenden Lehrkräfte einen Bewertungsbogen (Kapitel 4.2.5, Anhang A.4). Durch die Darbietung eines kriteriengeleiteten und Likert-skalierten Bewertungsbogens erhielten die angehenden Lehrkräfte nicht nur die Kriterien für ein gutes Lehr-Lern-Material, sondern konnten diese Kriterien auch durch die Ankreuzmöglichkeit selbst einschätzen.

Die Generierung und Darbietung der beiden Feedbacktypen intern und extern erfolgten in identischer Form und Weise, nur mit Unterscheidung des Feedbackgebenden. Dies ermöglichte einen direkten Vergleich des internen und externen Feedbacks. Beim internen Feedback wurde der Bewertungsbogen direkt von der angehenden Lehrkraft selbst ausgefüllt, die auch das Lehr-Lern-Material entwarf. Beim externen Feedback wurde der Bewertungsbogen für dieselbe angehende Lehrkraft von Seiten der Studienleitung ausgefüllt.

Während die Interventionsgruppe 1 sowohl mit systematischem internem und externem Feedback arbeitete, arbeitete die Interventionsgruppe 2 mit keinem systematischen externen Feedback und die Kontrollgruppe mit keinem systematischen Feedback. Um dennoch die Lern- und Arbeitszeit in allen drei Untersuchungsgruppen gleich zu halten, wurden die Zeiten in der Interventionsgruppe 2 und Kontrollgruppe, in denen kein systematisches externes Feedback stattfinden sollte, mit einem sogenannten Dummy-Feedback ersetzt (Kapitel 4.2.5).

In der Kontrollgruppe war zudem ein Ersatz für das systematische interne Feedback während der Phase der Feedbackgenerierung erforderlich. Die Testpersonen wurden dazu angeregt, ihr designtes Lehr-Lern-Material mit Schulbuchauszügen zu vergleichen, die zuvor passend zum Lernziel ausgewählt und bereitgestellt wurden. Durch diese Variation musste auch die Reflexionsphase der Kontrollgruppe angepasst werden. Dabei überlegten sie anhand des

Vergleichs zwischen ihrem eigenen Lehr-Lern-Material und den Schulbuchauszügen, wie sie die nächsten Lehr-Lern-Materialien besser gestalten können. Für den Vergleich gab es von Seiten der Studienleitung keine Vorgaben, sodass dieser frei und unsystematisch stattfand.

4.1.4 Entwicklung und Fachlicher Inhalt des Rohmaterials für das Lehr-Lern-Material

Die Rohmaterialien kamen jeweils in den Übungsphasen der Interventionssitzungen zum Einsatz. Bei der Auswahl des fachlichen bzw. inhaltlichen Themas der Lehr-Lern-Materialien wurden mehrere Aspekte berücksichtigt: Erstens sollte der chemische Inhalt der Lehr-Lern-Materialien möglichst einfach gehalten werden, damit der vom Material ausgehende Cognitive Load möglichst gering ist und sich die angehenden Lehrkräfte auf das Designen der multiplen externen Repräsentationen und die Umsetzung der (fach-)didaktischen sowie kognitionspsychologischen Aspekte fokussieren können. Gleichzeitig sollte das Thema Möglichkeiten bieten, beim Designen mehrere unterschiedliche Repräsentationen miteinander zu kombinieren und miteinander zu verknüpfen. Nach Abgleich der Kernlehrpläne der benannten Schulformen der Sekundarstufe I (MSB NRW, 2008, 2011a, 2011b, 2013) und Durchsicht von Schulbüchern (z. B. Asselborn et al., 2004; Asselborn et al., 2014; Eisner et al., 2009; Ernst & Becker, 2009; Gietz et al., 2013) wurde das Thema des einfachen Teilchenmodells nach Dalton (Brown et al., 2007) ausgewählt. Die Einführung des einfachen Teilchenmodells kann beispielsweise im Anfangsunterricht des Faches Chemie der Klasse 7 im Gymnasium erfolgen (MSW NRW, 2008). Nach Festlegung des übergeordneten Themas wurden inhaltlich drei sehr ähnliche, miteinander zusammenhängende und unterrichtsrelevante Themenaspekte ausgewählt und folgende Lernziele für die Übungsphasen der drei Interventionssitzungen formuliert:

- Interventionssitzung 1: Die Schüler:innen können den Aufbau von Stoffen und Materie am Beispiel des Kugelteilchenmodells beschreiben.
- Interventionssitzung 2: Die Schüler:innen können Aggregatzustände und ihre Übergänge mit Hilfe des Kugelteilchenmodells erklären.
- Interventionssitzung 3: Die Schüler:innen können Arten und Zusammensetzungen von Stoffgemischen mit Hilfe des Kugelteilchenmodells erklären.

Die Unterrichtsthemen und Lernziele wurden zusammenhängend und aufeinander abgestimmt gewählt, damit die Proband:innen sich in den einzelnen Sitzungen nicht neu eindenken müssen und der Cognitive Load niedrig gehalten werden kann. Darüber hinaus werden dadurch weitere Störfaktoren der Intervention reduziert, sodass die Ergebnisse der einzelnen Sitzungen auch eine höhere Vergleichbarkeit erbringen (z. B. konkretes Fachwissen der angehenden Lehrkräfte). Ein weiterer Vorteil ist, dass dadurch die Übungs- und Feedbackprozesse in der zweiten und dritten Sitzung routiniert werden und die angehenden Lehrkräfte einen noch stärkeren Fokus auf die (fach-)didaktischen Probleme in ihrem Lehr-Lern-Material richten können. Die drei Themen wurden zwar möglichst ähnlich gehalten, beziehen sich dennoch auf drei getrennte und aufeinander aufbauende Lerninhalte. Gleichzeitig wurde der Prozess des Designens so authentisch und gleichzeitig methodisch kontrollierbar und vergleichbar wie möglich gehalten.

Nach Vorgabe des Lernziels war ein Ausgangsmaterial erforderlich, welches angepasst und verändert werden sollte, sodass ein im Unterricht einsetzbares Arbeitsmaterial entsteht. Um die von jeder einzelnen Testperson äußerst individuell und detailreich designten Materialien vergleichbar zu halten und auch die Inhaltsvalidität der Intervention zu gewährleisten, wurden weitere Rahmenbedingungen vorgegeben: Die angehenden Lehrkräfte sollten nicht nur ein Arbeitsblatt zu einem Lernziel gestalten, sondern eine Lernaufgabe, welche aus einem Text und einer Abbildung besteht. Dadurch, dass es sich um eine Lernaufgabe handelte, beinhalteten die Lehr-Lern-Materialien relevante Abschnitte eines Lernprozesses (Ausgangsphänomen, Aufgabenstellung, Unterstützungsteil aus Text und Bild).

Um sowohl die Arbeitszeit für das Designen als auch die Recherchezeit nach geeigneten Ausgangsmaterialien zu minimieren, erhielten die angehenden Lehrkräfte für jede Interventionssitzung jeweils ein Rohmaterial im bearbeitbaren Word-Dateiformat (z. B. Anhang A.3). Dieses Rohmaterial enthielt zum einen Text, der alle relevanten Informationen zum vorgegebenen Lernziel beinhaltete, und zum anderen ein Bildmaterial, das in vier unterschiedlichen Varianten dargeboten wurde. Die in den Rohmaterialien enthaltenen Bild- und Textelemente sind in Anhang A.2 dokumentiert.

Als Beispiel sind in Abbildung 43 die bereitgestellten Bildmaterialien zur ersten Interventionssitzung dargestellt. Die gezeigten vier Varianten repräsentieren grundlegend denselben fachlichen Inhalt, unterscheiden sich jedoch in ihrer fachdidaktischen Aufbereitung. Je nach Aufbereitung wurden unterschiedliche und unterschiedlich viele typische Schüler:innenvorstellungen adressiert. So ergibt sich eine Rangfolge für die vier Varianten:

- Variante 1 (links oben): Die makroskopische Ebene (Becherglas) wird mit der submikroskopischen Ebene (Teilchen) vermischt abgebildet (die Teilchen sind direkt im Becherglas zu sehen). Die Moleküle sind als durchsichtige Kugeln mit einer Hintergrundfarbe zu sehen (Schüler:innenvorstellung: zwischen den Teilchen befindet sich weitere Materie). Das Becherglas und die Teilchen sind in ähnlichen Größenrelationen dargestellt.
- Variante 2 (rechts oben): Der Hintergrund ist entfernt und es sind Beschriftungen vorhanden. Die Ebenen werden immer noch vermischt.
- Variante 3 (links unten): Weitere Beschriftungen sind hinzugefügt. Das Teilchenmodell wird als submikroskopischer Ausschnitt aus der makroskopischen Ebene (Stoffgemisch aus Zucker und Wasser) abgebildet.
- Variante 4 (rechts unten): Die Färbung der Teilchen wurde verändert, so dass die Farben auf der makroskopischen Ebene (Zucker/weiß und Wasser/blau) nicht mehr mit den Teilchenfarben auf submikroskopischer Ebene identisch sind.

Nach ähnlichem Prinzip wurden die Text- und Bildmaterialien der zweiten und dritten Sitzung vorbereitet (Abbildung 45, Abbildung 46). In der zweiten Sitzung (Aggregatzustände) wurde die Darbietung von vier Bildvarianten beibehalten. Zusätzlich wurden in der zweiten Sitzung Variationen im Rohtext eingebaut. Die angehenden Lehrkräfte sollten dadurch angeregt werden, aus mehreren Textfragmenten einen ganzheitlichen, zusammenhängenden, für die Lernenden verständlichen und zugleich fachlich und fachsprachlich korrekten Text zu formulieren. Formulierungen wie „beim Erwärmen wackeln die Teilchen hin und her“ oder „Eis ist wässrig bei 0 °C“ sollten den angehenden Lehrkräften als optimierbar erscheinen.

In der dritten Sitzung wurden durch den inhaltlichen Umfang des Lernziels die Varianten der Abbildung komplexer (Abbildung 47, Abbildung 48). Hierbei

ergab sich die Möglichkeit, nicht nur die Bilder der Ebenen und die Teilchenabbildungen zu variieren, sondern auch die Textelemente innerhalb der Abbildungen. Aufgrund der Komplexität der Abbildungen wurden lediglich drei Versionen dargeboten. Auch der Rohtext war in der dritten Sitzung umfangreicher als in den vorigen Sitzungen.

Diese Aufgabe zu bewältigen erforderte von den angehenden Lehrkräften nicht nur fachliches, (fach-)didaktisches und kognitionspsychologisches Wissen, sondern auch technische Kompetenzen. Beispielsweise war es wichtig, dass die angehenden Lehrkräfte mit ihrem Textbearbeitungstool (z. B. Microsoft Word, Apache OpenOffice, Apple Pages) so kompetent umgehen können, dass sie ihre Ideen und Vorstellungen bzgl. ihres Arbeitsblattes umsetzen können. Um auch diese Kompetenzen zu erfassen und zu kontrollieren, wurden daher im Prätest zu den Kontrollvariablen (Tabelle 4) geeignete Items hinzugefügt, die die Erfahrungen der angehenden Lehrkräfte in der Gestaltung von Lehr-Lern-Materialien abfragen. Falls es dennoch Aspekte oder Ideen gab, die die angehenden Lehrkräfte in ihren Arbeitsblättern nicht umsetzen konnten, hatten sie bei der Abgabe der finalisierten Arbeitsblätter die Gelegenheit, ihre Ideen mitzuteilen.

4.2 Testinstrumente und Pilotierungen

Im Folgenden wird dargelegt, mit welchen Testinstrumenten die abhängigen Variablen und Kontrollvariablen erhoben wurden. Die Erhebungsinstrumente der Prä- und Posttests wurden überwiegend aus bestehenden Tests adaptiert. Auch die Tests zum repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen sind aus bestehenden Instrumenten hervorgegangen, wurden inhaltlich jedoch stark verändert, sodass eine Pilotierung erfolgte. Die Instrumente, die während der Intervention zum Einsatz kamen (z. B. der Bewertungsbogen), wurden selbst entwickelt.

4.2.1 *Demografische Daten und persönliche Ressourcen*

Zur Erhebung der personenbezogenen Kontrollvariablen gelten die „Skalen zur Qualität der Lehrerbildung“ (Abs et al., 2005) als Grundlage. Das Skalenhandbuch beinhaltet u. a. Items zur Befragung angehender Lehrkräfte im Vorbereitungsdienst zu ihrer beruflichen Situation und zur Person, woraus mehrere (Sub-)Skalen entnommen und teilweise mit eigenen Variablen ergänzt wurden. Folgende Variablen wurden zunächst für die *demografischen Daten* eingesetzt:

- ZfsL (Eigenentwicklung)
- Lehramtsform (Eigenentwicklung)
- Schulform (Abs et al., 2005)
- Studiengang, durch den die Zulassung in den Vorbereitungsdienst erworben wurde (Eigenentwicklung und berichtet als Dienstgrad)
- Absolvierte Zeit im Vorbereitungsdienst (Eigenentwicklung)
- Dauer einer vorigen Vertretungsstelle (Eigenentwicklung)
- Alter (Abs et al., 2005)
- Abitur-Note (Abs et al., 2005)
- Hochschulabschluss-Note (angepasst aus: Examensnote; Abs et al., 2005)
- Bundesland, in dem der Hochschulabschluss erworben wurde (Eigenentwicklung)
- Muttersprache(n) (nach Grosjean, 2021, adaptiert aus: Muttersprache; Abs et al., 2005)
- Geschlecht (Abs et al., 2005)

Die Variablen wurden so ausgewählt, dass die Stichprobe ausführlich beschrieben werden kann und zudem Varianz- und Zusammenhangsanalysen bezogen auf die entsprechenden Variablen berechnet werden können. Alle Items sind im Multiple-Choice-Format oder mit einem offenen Eingabefeld angelegt. Die einzelnen zugehörigen Ausprägungsformen bzw. Antwortmöglichkeiten sowie Kennwerte sind im Ergebnisteil (Kapitel 5.1.1) mit ihren Häufigkeiten einsehbar.

Die *persönlichen Ressourcen* der angehenden Lehrkräfte wurden auf Grundlage einer Selbsteinschätzung erfasst und beziehen sich auf das Bewältigen von Aufgaben und Problemen im schulpraktischen Handeln. Sie umfassen die Skalen *Belastungen und Bewältigungsstrategien* sowie *Professionelles Selbstverständnis*, welche wiederum aus Subskalen bestehen. Die zugehörigen Items und Antwortmöglichkeiten wurden aus dem Skalenhandbuch aus dem Abschnitt „Gestaltbare individuelle Bedingungen für Referendarinnen und Referendare“ (Abs et al., 2005) unverändert übernommen. Zur besseren Nachvollziehbarkeit und Prägnanz wurden jedoch die Bezeichnungen der originalen Skalenbezeichnungen angepasst. Die Gruppierung der aufgelisteten Subskalen in die beiden Überkategorien (Belastungen und Bewältigungsstrategien, Professionelles Selbstverständnis) erfolgte auf inhaltlicher Basis für eine höhere sachliche Nachvollziehbarkeit der Ergebnisse. In die Überkategorie *Belastungen und Bewältigungsstrategien* wurden folgende Subskalen einbezogen:

- Zeitdruck (umbenannt aus: Zeit fürs Referendariat)
- Berufliche Belastungen (umbenannt aus: Belastungen im Beruf)
- Lernstrategien (umbenannt aus: Lernstrategien, Zusammenhänge erkennen)
- Elaborationsmuster (umbenannt aus: Muster in der Bewältigung von Aufgaben und Belastungen)
- Soziale Ressourcen (umbenannt aus: Soziale Ressourcenorientierung in der Bewältigung von Aufgaben; Abs et al., 2005)

Aus der Überkategorie *professionelles Selbstverständnis* wurden diejenigen Subskalen ausgewählt, die eine hohe Passung zum Forschungsvorhaben aufweisen. Auf diese Weise wurde die inhaltliche Validität der Überkategorie erhöht. Um die Selbstwirksamkeitserwartung von Lehrkräften der naturwissenschaftli-

chen Unterrichtsfächer zu erheben, wurde ergänzend die Skala zur Selbstwirksamkeit nach Bleicher (2004) eingesetzt. Insgesamt umfasst die Überkategorie zum professionellen Selbstverständnis folgende Subskalen:

- Einstellungen zur Wichtigkeit der Wissenschaft (umbenannt aus: Einstellungen zum Verhältnis von Wissenschaft und Praxis)
- Selbstwirksamkeitserwartung an sich selbst (übersetzt aus: Personal Science Teaching Efficacy Belief; Bleicher, 2004)
- Selbstwirksamkeitserwartung an den eigenen Chemieunterricht (übersetzt aus: Science Teaching Outcome Expectancy; Bleicher, 2004)

Um die bisherigen Erfahrungen der Testpersonen in der *Gestaltung von Lehr-Lern-Materialien* zu erfassen, wurden erneut Items aus den „Standards in der Lehrer/-innenbildung“ (Abs et al., 2005, S. 59) eingesetzt. Die Subskalen wurden so umbenannt und angepasst, dass folgende Teilkompetenzen des Gestaltens von Lehr-Lern-Materialien erfragt wurden:

- Anknüpfen an das Vorwissen von Schüler:innen
- Vorbereiten von Unterrichtsmaterial
- Logisches Gliedern ausgewählter Lerninhalte
- Strukturieren chemischer Inhalte mit Hilfe des Lehrplans über mehrere Klassenstufen
- Auswählen von Unterrichtsmaterial auf Basis eigener Bewertungen

Dabei gaben die Testpersonen an, in welcher Ausbildungssituation die Teilkompetenzen erlernt wurden (Multiple Choice, Abbildung 9) und wie sicher sie sich in der jeweiligen Kompetenz einschätzen (Ranking zwischen 0-4, Tabelle 25). Zum besseren Verständnis der genannten Variablen, (Sub-)Skalen und Items ist eine Einsicht in den Ergebnisteil (Kapitel 5.1) empfehlenswert, da hier alle Itembeschreibungen, Formate und Auswahlmöglichkeiten ersichtlich werden.

4.2.2 Allgemeines Fachwissen und allgemeines fachdidaktisches Wissen

Aus dem Forschungsprojekt ProwiN (Professionswissen von Lehrkräften in den Naturwissenschaften) existieren erprobte Tests zum fachlichen und fachdidaktischen Wissen (Strübe, 2021). Der Fachwissenstest besteht aus 29 Items im Single-Choice-Format und ist für eine Bearbeitungszeit von 25 Minuten ausgelegt (Strübe, 2021). Er wurde an einer Stichprobengröße von $n = 355$ Lehrkräften

erprobt und weist eine Reliabilität von $\alpha = 87$. auf. Um die Gesamtbearbeitungszeit des Prätests von 120 Minuten einhalten zu können, wurde der Fachwissens-test auf 18 Items gekürzt, sodass er in 15 Minuten bearbeitbar ist. Zur Kürzung des Tests wurden einerseits die Kennwerte (z. B. Aufgabenschwierigkeit, Item-trennschärfe) aus dem verfügbaren ProwiN-Datensatz (Dollny, 2011) analysiert und andererseits darauf geachtet, dass die Inhaltsvalidität gewahrt bleibt (Ta-belle 7). Während die Erhebungen des ProwiN-Projekts im Paper-Pencil-Format stattfanden, erfolgte für den vorliegenden Einsatz eine Übertragung der Aufga-ben in ein Online-Format (Moodle).

Zunächst wurden die Items auf der inhaltlichen Ebene (Tabelle 7) betrachtet: Bei der Testentwicklung im ProwiN-Projekt wurden die Items nach Inhaltsfel-dern (Atombau, Chemische Bindung, Säuren und Basen) und ihrem Anspruchs-grad (Sekundarstufe I, II, Grundstudium) unterteilt (Dollny, 2011). Da es sich bei den Testpersonen in der vorliegenden Untersuchung um angehende Lehr-kräfte aller Schulformen handelt und nicht alle mit den Inhalten der Sekundar-stufe II vertraut sind, wurden zunächst die Items des Niveaus der Sekundarstufe II gestrichen. Außerdem gehört auch der in der Intervention behandelte Fachin-halt (Einführung des Teilchenmodells) in die Sekundarstufe I, sodass die Ober-stufeninhalte während der Intervention vollkommen unberührt bleiben. Im nächsten Schritt wurde überprüft, in welchen Themenbereichen bzw. Feldern es eine Mehrzahl an Items gibt. Aus diesen wurde dann nach Möglichkeit das Item mit den bestmöglichen Kennwerten übernommen. Berechnet mit dem originalen ProwiN-Datensatz ergab sich dadurch ein Test aus 18 Items mit einer hohen Reliabilität von $\alpha = .79$.

Gleichermaßen erfolgte die Anpassung des fachdidaktischen Testteils. Hier wurden Rating-Scale-Aufgaben eingesetzt: Im Aufgabenstamm wird eine unter-richtliche Situation mit einem fachdidaktischen Problem geschildert (z. B. Ver-besserung einer chemischen Reaktionsgleichung, die eine Referendarin an die Tafel schrieb). Anschließend werden vier fachdidaktische Antwortalternativen aufgelistet, die von den Testpersonen auf einer sechsstufigen Notenskala gerankt werden sollen. Somit gibt es keine dichotom richtigen oder falschen Alternati-ven. Bei der Auswertung der Antworten wird lediglich auf das Ranking der Ant-wortalternativen geachtet (Dollny, 2011). Dieser Testteil weist eine Reliabilität

von $\alpha = .81$ auf und besteht ursprünglich aus 17 Items, die in einer Bearbeitungszeit von 20 Minuten gelöst werden sollen (Strübe, 2021). Der Test ist so konzipiert, dass nicht curriculare Inhalte, sondern fachdidaktische Themen wie Experimente, Modelle und Schüler:innenvorstellungen thematisiert werden (Dollny, 2011; Tabelle 8). Um auch diesen Test in 15 Minuten bearbeiten zu können, wurde der fachdidaktische Teil nach dem gleichen Vorgehen wie im Falle des Fachwissenstests auf 13 Items gekürzt. Mit dem entsprechend gekürzten ProwiN-Datensatz (Dollny, 2011) ergibt sich eine Reliabilität von $\alpha = .76$.

Tabelle 7: Inhaltliche Verteilung der Testitems zum allgemeinen Fachwissen

Curriculare Zuordnung	Wissensart	Atombau und Periodensystem	Chemische Bindungen	Säuren und Basen
Sekundarstufe I	Inhaltswissen	DAS3	DBS1	DCS1, DCS2, DCS3
	Handlungswissen	-	PBS1	PCS1
	Begründungswissen	KAS1	KBS1	KCS1
Sekundarstufe II	Inhaltswissen	DAO1	-	DCO1, DCO2
	Handlungswissen	PAO1	-	PCO1, PCO2, PCO3
	Begründungswissen	KAO1	KBO1	KCO1
Grundstudium	Inhaltswissen	DAG1	DBG1	DCG1
	Handlungswissen	PAG1	PBG1	PCG1
	Begründungswissen	KAG2	KBG1	KCG1

Tabelle 8: Inhaltliche Verteilung der Testitems zum allgemeinen fachdidaktischen Wissen

Fach- didakt. Facetten	Wissensart	Atombau u. Perioden- system	Chem. Bindun- gen	Säuren und Basen	Themen- unspezi- fisch
Experi- mente	Inhalts- wissen	-	-	DCE1	-
	Handlungs- wissen	PAE1	-	-	PAIE2
	Bedingungs- wissen		-	KCE1	KAIE1
Modelle	Inhaltswissen	-	-	-	DAIM2, DAIM3
	Handlungs- wissen	-	-	-	-
	Bedingungs- wissen	-	KBM1	-	-
Schüler:in- nenvor- stellungen	Inhaltswissen	DAS1, DAS2	DBS1	-	PAS1
	Handlungs- wissen	-	-	PCS1	-
	Bedingungs- wissen	KAS1	KBS1 , KBS2	KCS1	-

4.2.3 Repräsentationsbezogenes Fachwissen und fachdidaktisches Wissen

Auch zum repräsentationsbezogenen Wissen angehender Lehrkräfte kann auf vorhandene Testinstrumente zurückgegriffen werden. So wurden im Rahmen des KiL-Projekts (Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen) Tests entwickelt, die das repräsentationsbezogene Fachwissen sowie das fachdidaktische Wissen messen (Kleickmann et al., 2014; Taskin et al., 2015b; Taskin et al., 2012). Beide Testteile beziehen sich jedoch auf Repräsentationen der organischen Chemie und bilden die Fachinhalte der vorliegenden Intervention eher geringfügig ab. Daher wurden beide Teile zunächst inhaltlich analysiert und anschließend adaptiert und pilotiert.

Der ursprüngliche Fachwissenstest von Taskin et al. (2012) besteht aus 36 Items im geschlossenen oder halboffenen Aufgabenformat. Berechnungen aus Daten von $n = 220$ Lehramtsstudierenden liefern Reliabilitätswerte von $\alpha = .84$

(Kleickmann et al., 2014). Inhaltlich bezieht sich der Test auf Repräsentationen, welche vorwiegend in der organischen Chemie genutzt werden (z. B. Sesselschreibweise, Newman-Projektionsformel, Tautometrie, Energieniveaudia-gramme), die zudem eher das universitäre Niveau abdecken.

Ein ähnliches Problem, welches die Inhaltsvalidität des Tests einschränkt, lag auch für den Testteil zum repräsentationsbezogenen fachdidaktischen Wissen vor: Der Test besteht aus 25 Items im halboffenen oder offenen Aufgabenformat. Es werden Reliabilitätswerte von $\alpha = .67-.77$ berichtet (Kleickmann et al., 2014). Die Items sind so konstruiert, dass der Itemstamm eine Repräsentation aus einer unterrichtlichen Situation wiedergibt (z. B. fachlich falsches Kugelteilchenbild vom Wasser-Molekül) und im Anschluss die möglichen Fehlinterpretationen und Denkfehler erfragt, die Lernende mit Blick auf das Beispiel haben können (z. B. Sauerstoff- und Wasserstoffatome sind gleich groß, die Atome im Molekül sind falsch angeordnet). Daneben gibt es einen großen inhaltlichen Schwerpunkt in der organischen Chemie (z. B. Diamantstruktur) oder auch allgemeine und zu offene Aufgaben (z. B. allgemeine Beschreibung der Struktur-Eigenschaftsbeziehung, Nennen von Fördermöglichkeiten, Fragen zum Spiralcurriculum; Taskin & Bernholt, 2012).

Ein Ziel der Adaption der beiden Testteile war die Erhöhung der Inhaltsvalidität. Um dies systematisch anzugehen, wurde nach dem Vorbild des ProwiN-Projekts (Tepner et al., 2012) ein Testmodell entwickelt. Ziel eines solchen Modells war die Definition relevanter Aufgabenmerkmale, die das repräsentationsbezogene Fachwissen und fachdidaktische Wissen ausdrücken. Im Zuge der Operationalisierung können vorhandene Items darin eingeordnet werden, um dadurch inhaltliche Lücken der bestehenden Testteile aufzudecken. Das Modell umfasst drei Dimensionen (Abbildung 6): (1) die Repräsentationsformen, die im Chemieunterricht eingesetzt werden (z. B. Leisen, (1998, 2005), (2) die chemischen Inhaltsfelder der vier klassischen Disziplinen der Chemie (Brown et al., 2007) und (3) die Kategorien repräsentationsbezogenen Wissens nach Nitz (2012).

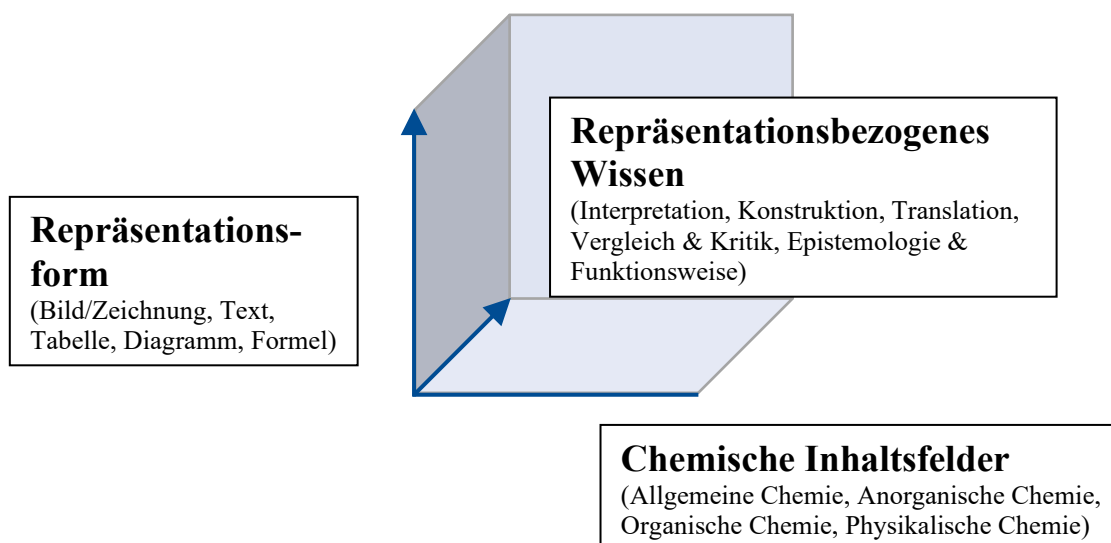


Abbildung 6: Testmodell zur Konstruktion der Testitems zum repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen

Um eine inhaltliche Einordnung der bereits bestehenden Testitems in das Testmodell durchführen zu können, wurde zudem ein Raster erstellt (Tabelle 9). Die Einordnung der Testitems (1.1-1.23; Taskin et al., 2015a) zeigt, dass sich die meisten Items auf chemische Formeln und Texte beziehen. Aus den bestehenden Items wurden zunächst inhaltlich passende Items mit möglichst guten Kennwerten (Taskin et al., 2015a) ausgewählt. Items, die inhaltlich ein repräsentationsbezogenes Wissen auf universitärem statt schulischem Niveau abfragen, wurden ausgeschlossen (z. B. Item 1.22, Enantiomerie). Als nächstes wurden neue Items zu den leeren Feldern des Testmodells entwickelt. Hierbei wurde darauf geachtet, dass es sich bei den Iteminhalten um curricular valide und authentische Repräsentationen handelt. Der so entwickelte fachdidaktische Testteil für die Pilotierung umfasst 17 Items und sollte in 25 Minuten bearbeitet werden.

Tabelle 9: Inhaltliche Verteilung der Testitems zum repräsentationsbezogenen Fachwissen und repräsentationsbezogenen fachdidaktischen Wissen vor der Pilotierung und zugehörige Kürzungen

Repräsen- tations- form	Kategorien der Repräsentationskompetenz				Epistemo- logie und Funktions- weise
	Interpretation	Konstruktion	Translation	Vergleich und Kritik	
Bild, Foto- grafie	-	-	RTN04 (AC)	RVN02 2.14	REN05 REN06 REN11
Strich-/ Schema- zeichnung	-	-	RTK01/1.2 (NC) RTN01 (AC)	RVK02/1.6 (AC) RVN04 2.7 2.8 2.9 2.21 2.22 2.23 2.24 2.25	REK01/2.1 REK02/2.3 REK03/2.5 REK04/2.6 REN12
Technische Zeichnung/ Versuchs- aufbau	RIN04 (PC)	-	RTN03 (AC)	RVN05 2.10 2.11 2.12 2.13	REN13 2.16 2.17 2.19
Text	1.7 (PC) 1.11 (AC) 1.12 (AC) 1.13 (NC)	1.14 (AC)	1.8 (OC) 1.9 (NC) 1.10 (OC)	2.18 2.20	REN09
Strukturie- rendes Diagramm	RIN01 (AC)	-	RTN02 (PC)	-	-
Tabelle	-	-	-	-	REN07
Quantifizie- rendes Diagramm	RIN05 (PC) RIN03 (PC)	RKN01 (AC)	-	RVN06	REN14
Chemische Formel	RIK01/1.18 (AC) 1.24 (OC) 1.26 (OC) 1.28 (OC)	RKK04/1.4 (NC) RKK02/1.17 (OC) RKK03/1.21 (OC) RKK05/1.23 (OC) 1.27 (OC) 1.30 (OC) 1.34 (OC)	1.15 (OC) 1.16 (OC) RTK02/1.19 (OC) RTK03/1.20 (OC) 1.25 (OC) 1.29 (OC) 1.31 (OC) 1.32 (OC) 1.33 (OC)	RVK01/1.1 (AC) 1.5 1.22 (OC) RVN01 (OC) 2.2 2.5 2.6 RVN03	2.4
Mathemati- sche Formel	-	-	-	-	-

Anmerkung. Schwarz: Items zum repräsentationsbezogenen Fachwissen, *grau und kursiv*: Items zum repräsentationsbezogenen fachdidaktischen Wissen.

Beide Testteile wurden gemeinsam mit dem Test zu den repräsentationsbezogenen Überzeugungen (Kapitel 4.2.4) pilotiert. Ziel der Pilotierung war die Erprobung der adaptierten Tests und die Ermittlung der Qualität bzw. Kennwerte. Hierzu fand eine Querschnittserhebung im Paper-Pencil-Format mit $N = 50$ Lehramtsstudierenden aus Masterstudiengängen (58 % weiblich) von drei Universitäten in Nordrhein-Westfalen statt. Die Studierenden befanden sich mehrheitlich im ersten (30%) und zweiten Semester (46%) bzw. vor oder im Praxissemester. Aus den Auswertungen der Testergebnisse gingen mehrere Kennwerte hervor (Tabelle 10). Um die Reliabilitäten zu erhöhen und die Itemtrennschärfen zu verbessern, wurden die Tests auf 14 bzw. 12 Items gekürzt. Dies erfolgte durch die nachgelagerte Analyse der Kennwerte der einzelnen Items, welche im Detail im Anhang zu finden sind (Tabelle 51, Tabelle 52).

Tabelle 10: Testkennwerte aus der Pilotierung zum repräsentationsbezogenen Fachwissen und repräsentationsbezogenen fachdidaktischen Wissen

Test	Item-anzahl	Scores [M , SD]	Reliabilität (α)	Itemtrennschärfe (r_{it})	Popularitätsindex (p)
Rep. FW	12	52.49, 14.07	.75	.12-.61	.14-.94
Rep. FDW	12	49.35, 14.40	.63	.13-.48	.06-.84

Anmerkung. Es handelt sich um relative Summenscores (vgl. Kapitel 4.3.1).

Insgesamt verfügen die Tests in der Pilotierung über noch akzeptable Kennwerte (nach Blanz, 2021). Zu berücksichtigen ist, dass die originalen Tests aus dem KiL-Projekt über ähnliche und nur leicht höhere interne Konsistenzen verfügen (rep. FW: $\alpha = .83$ -.84, rep. FDW: $\alpha = .67$ -.77; Kleickmann et al., 2014). Ebenso ist zu beachten, dass vor allem zum repräsentationsbezogenen fachdidaktischen Wissen zahlreiche neue selbstkonstruierte Items hinzugefügt und neu erprobt wurden. Die Erwartung war hier daher, dass sich die Kennwerte in der Hauptstudie verbessern werden. Erster Grund für die Annahme ist die doppelt so große Stichprobe in der Hauptstudie und der zweite Grund die Korrektur bzw. Opti-

mierung einiger Items. Beispielsweise wurden Items mit zwei inhaltlichen Abfragen (z. B. „Beschriften Sie den *oberen* und *unteren* Pfeil im Diagramm.“) in zwei Items geteilt. Des Weiteren gab es einige Items im halboffenen Aufgabenformat, bei der die Testpersonen ihre eigenen freien Antworten eintragen sollten (z. B. „Entnehmen Sie aus dem Diagramm zum Born-Haber-Kreisprozess die fehlende chemische Reaktionsgleichung und stellen Sie diese korrekt auf.“). Hier wurden die häufigsten drei falschen Antworten der Testpersonen genutzt, um eine Single-Choice-Aufgabe mit drei Distraktoren zu konstruieren. Das Aufgabenformat wurde somit von halboffen zu geschlossen angepasst. Die gesamten Optimierungen und inhaltlichen Zuordnungen sind in der Tabelle 11 sowie im Anhang (Tabelle 51, Tabelle 52) zu finden.

Vor Beginn der Hauptstudie mussten die Tests zum repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen angepasst werden. Grund hierfür ist, dass die Hauptstudie auf Basis von Absprachen mit den Fachleitungen und ZfsL so geplant war, dass sie im Zeitraum von Mai bis Juli 2020 durchgeführt werden sollte. Da sich dieser Zeitraum mit den pandemisch bedingten Schulschließungen, Lockdowns etc. überschneidet, waren Testungen und Interventionen, welche ursprünglich in Präsenz stattfinden sollten, nicht mehr möglich. So mussten die Testitems in ein Online-Format überführt werden. Dadurch konnten einige wenige Items in ihrer ursprünglichen Variante (z. B. Zeichnen von Strukturformeln, Eintragen von Elektronenverschiebungspfeilen) nicht eingesetzt werden und wurden komplett ausgelassen. Beide Testteile wurden so reduziert und finalisiert, dass sie in jeweils 15 Minuten bearbeitet werden konnten. Die aus den Ergebnissen berechneten Kennwerte sind im Ergebnisteil zur Hauptstudie (Kapitel 5.2) sowie im Anhang C dokumentiert.

Tabelle 11: Inhaltliche Verteilung der Testitems zum repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen nach der Pilotierung und zugehörige Kürzungen

Repräsen- tations- form	Kategorien der Repräsentationskompetenz				
	Interpretation	Konstruktion	Translation	Vergleich und Kritik	Epistemo- logie und Funktions- weise
Bild, Foto- grafie	-	-	RTN04	<i>RVN02</i>	<i>REN05</i> <i>REN06</i> <i>REN11</i>
Strich-/ Schema- zeichnung	-	-	RTK01 RTN01	RVK02 <i>RVN04</i>	<i>REK01</i> <i>REK01a</i> <i>REK01b</i> <i>REK02</i> <i>REK03</i> <i>REK04</i> <i>REN12</i>
Technische Zeichnung/ Versuchs- aufbau	RIN04	-	<i>RTN03</i>	<i>RVN05</i>	<i>REN13</i>
Text	-	-	-	-	<i>REN09</i>
Strukturie- rendes Diagramm	<i>RIN01</i> RIN01a RIN01b	-	<i>RTN02</i> RTN02a RTN02b	-	-
Tabelle	-	-	-	-	<i>REN07</i>
Quantifizie- rendes Diagramm	RIN05 <i>RIN03</i>	<i>RKN01</i>	-	<i>RVN06</i>	<i>REN14</i>
Chemische Formel	<i>RHK01</i>	<i>RKK04</i> <i>RKK02</i> <i>RKK03</i> RKK05	RTK02 RTK03	RVK01 <i>RVN01</i> <i>RVN03</i>	-
Mathemati- sche Formel	-	-	-	-	-

Anmerkung. Schwarz: Items zum rep. Fachwissen, orange: Items zum rep. fachdidaktisches Wissen.

4.2.4 Repräsentationsbezogene Überzeugungen

Als Grundlage für die Erfassung der repräsentationsbezogenen Überzeugungen dient der Fragebogen von Nitz et al. (2011), welcher sich mit der Relevanz von Fachsprache im naturwissenschaftlichen Unterricht befasst. Erprobt wurde der Fragebogen an Studierenden des gymnasialen Lehramts ($n = 33$) und angehenden Lehrkräften im Vorbereitungsdienst ($n = 33$) mit mindestens einem naturwissenschaftlichen Unterrichtsfach. Der Fragebogen besteht aus 20 Items, deren Aussagen auf einer vierstufigen Ratingskala beurteilt werden. Die drei ermittelten Skalen (1) „Vermittlung von Fachsprache als Ziel des naturwissenschaftlichen Unterrichts“, (2) „Verständlichkeit des naturwissenschaftlichen Unterrichts“ und (3) „Fachsprache als Qualitätsmerkmal des naturwissenschaftlichen Unterrichts“ bestehen jeweils aus drei bis fünf Items und weisen eine interne Konsistenz von $\alpha = .66$ – $.76$ auf (Nitz et al., 2011).

Die erste Adaptierung und qualitative Erprobung des Erhebungsinstruments via Stimulated Recall Interviews fand im Rahmen einer Masterarbeit statt (Peukert, 2018). Die quantitative Erprobung zur Analyse der Testkennwerte fand im Zuge der Pilotierung der beiden Tests zum repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen statt. Die Kennwerte zur Reliabilität und Trennschärfe waren mit einem Cronbachs Alpha von $\alpha = .62$ bei 17 Items nicht zufriedenstellend (Tabelle 12).

Tabelle 12: Testkennwerte aus der Pilotierung zu den repräsentationsbezogenen Überzeugungen

Test	Item- anzahl	Rating [M , SD]	Reliabili- tät (α)	Itemtrenn- schärfe (r_{it})	Populari- tätsindex (p)
Rep. Überzeu- gungen	17	1.87, .63	.62	-.06-.56	.29-.89

Anmerkung. Skala: 0 = „sehr geringe Relevanz von Repräsentationen“ bis 3 = „sehr hohe Relevanz von Repräsentationen“.

So wurde der gesamte Fragebogen vor dem Einsatz in der Hauptuntersuchung überarbeitet. Ausgehend von der Arbeit von Peukert (2018) wurden die Skalen nach der vorliegenden Pilotierung ergänzt:

1. Einsatz von Repräsentationen als Ziel des Chemieunterrichts (UBZI),
2. Verständlichkeit des Chemieunterrichts durch fachtypische Repräsentationen (UBVE),
3. Verständlichkeit des Chemieunterrichts durch die didaktisch angemessene Auswahl und Gestaltung der Repräsentationen (UBGE),
4. Repräsentationen als Qualitätsmerkmal des Chemieunterrichts (UBQU).

Die in der Hauptstudie eingesetzten Items sind im Anhang A.5 aufgelistet. Grundsätzlich wurde bei der Überarbeitung der Items darauf geachtet, dass sie die Testpersonen so direkt wie möglich ansprechen (z. B. Ich-Formulierungen und Konkretisierungen wie „Als Chemielehrkraft sollte ich...“, „Wenn ich für meinen Unterricht plane...“). So wurde auch der allgemeine naturwissenschaftliche Bezug (z. B. „im naturwissenschaftlichen Unterricht“) auf das Unterrichtsfach Chemie spezialisiert (z. B. „im Chemieunterricht“). Des Weiteren wurden in den Items genauere Fallbeispiele genannt und diese mit Beispielen ergänzt (z. B. „Im Chemieunterricht sollten Schüler:innen lernen, Inhalte aus einer Darstellungsform (z. B. Tabelle) in eine andere Darstellungsform (z. B. Liniendiagramm) umzuwandeln“). Für die neue Testversion konnte in der Hauptuntersuchung eine interne Konsistenz von $\alpha = .81$ erreicht werden (Kapitel 5.3, Anhang Tabelle 56).

4.2.5 *Repräsentationsbezogene Qualität der Lehr-Lern-Materialien*

Das Testinstrument zur Bewertung der repräsentationsbezogenen Qualität von Lehr-Lern-Materialien ist ein Tool für angehende Lehrkräfte, um die Qualität ihrer Unterrichtsmaterialien einzuschätzen. Bezogen auf den Einsatz während der Intervention (Tabelle 5) erfüllt der Bewertungsbogen zwei zentrale Aufgaben: Die angehenden Lehrkräfte designen ein Lehr-Lern-Material (Phase: Übung) und erhalten den Bewertungsbogen, um ihr Lehr-Lern-Material zu bewerten (internes Feedback, Phase: Feedbackgenerierung). Anschließend erhalten sie den gleichen Bewertungsbogen individuell ausgefüllt von der Studienleitung (externes Feedback, Phase: Reflexion). Zuletzt vergleichen sie beide Feedbacks miteinander und reflektieren darüber, welche Verbesserungen sie bei der nächsten Materialerstellung vornehmen können. Dieser Ablauf findet vollständig in der Interventionsgruppe 1 statt, während die Phasen der Feedbackgenerierung und/oder Reflexion in der Interventionsgruppe 2 und Kontrollgruppe leicht abgewandelt stattfinden (Kapitel 4.1.3).

Auf diese Weise werden den angehenden Lehrkräften alle drei Komponenten eines vollständigen und effektiven Feedbacks dargeboten (Hattie & Timperley, 2007). Das Feed Up, welches das Lernziel angibt, wird mit dem Bewertungsbogen detailliert vermittelt. Die angehenden Lehrkräfte sollen in der Lage sein, Lehr-Lern-Materialien zu designen, die über eine möglichst hohe Qualität verfügen. Die einzelnen Kriterien, die die Gesamtqualität ergeben, sind im Bewertungsbogen aufgeführt. Mithilfe der Skala lässt sich zudem einschätzen, auf welchem Niveau sich die Lehrkräfte aktuell befinden und wie groß das Verbesserungspotenzial ist. Somit ist auch das Feed Back bzw. die Einschätzung des Ist-Werts und der Vergleich mit dem Soll-Wert sichergestellt. Zuletzt erfolgt das Feed Forward, indem nächste Schritte geplant werden. Da die Reflexion bei jeder Testperson individuell verläuft, zumal die einzelnen Lehr-Lern-Materialien bereits individuell designt sind, werden die Denk- und Arbeitsergebnisse dieser Phase von den angehenden Lehrkräften in freier Form als Notizen festgehalten. Es folgt hier lediglich die Analyse durch die quantitativen Daten des Bewertungsbogens, ob die zukünftigen Lehr-Lern-Materialien an Qualität gewinnen. Da die Spannweite an Unterrichts- und Arbeitsmaterialien sehr groß ist, wurde das Testinstrument möglichst universell konzipiert. Der Bewertungsbogen fokussiert Arbeitsblätter, die eine Lernaufgabe beinhalten und in ihrem informativen Teil aus einer Text-Bild-Kombination bestehen. Der Bewertungsbogen wurde in drei Etappen entwickelt, bevor er in der Hauptstudie eingesetzt wurde. Dies erfolgte im Rahmen dreier aufeinander aufbauenden Masterarbeiten (Mühlhaus, 2020; Tonyali, 2018; Ude, 2016): Im ersten Entwicklungsschritt wurde auf inhaltlicher Ebene gearbeitet. Hierzu führte Ude (2016) u. a. eine ausführliche Literaturrecherche durch und kategorisierte alle bekannten Repräsentationsformen, die in den Naturwissenschaften und im naturwissenschaftlichen Unterricht vorkommen und benannte ihre fachspezifischen Eigenschaften. Im nächsten Schritt wurde das Kategoriensystem inhaltlich auf multiple Repräsentationen und den Chemieunterricht angepasst und in das Format eines Bewertungsbogens überführt (Tonyali, 2018). Im letzten Schritt wurde der Bewertungsbogen, der bisher auf das Paper-Pencil-Format ausgerichtet war, in ein Moodle-basiertes digitales Tool umgewandelt. Zudem wurde der Bewertungsbogen erprobt (Mühlhaus, 2020).

Angelehnt an die im Theorieteil zusammengefassten Grundlagen über die Eigenschaften multipler externer Repräsentationen (Kapitel 2.1.3) wurde für den Bewertungsbogen zunächst ein Kriterienkatalog zusammengestellt. Wie im Interventionsdesign beabsichtigt, wurde der Bewertungsbogen so konzipiert, dass er auf multiple externe Repräsentationen in Lehr-Lern-Materialien anwendbar ist. Er beginnt mit der Klassifizierung der vorliegenden Repräsentation(en), so dass im Anschluss überprüft werden kann, inwiefern die Kriterien für die jeweiligen Repräsentationsformen und -ebenen zutreffen. Angepasst an die Intervention wurden die Kriterien so zusammengestellt, dass sich etwa 100 Kriterien auf insgesamt 22 Kategorien verteilen (Anhang A.4).

Bezogen auf die Skalierung wurden zunächst verschiedene Skalenbreiten untersucht, bevor eine fünfstufige Skala (1 = trifft vollständig zu, 2 = trifft eher zu, 3 = trifft teilweise zu, 4 = trifft eher nicht zu, 5 = trifft gar nicht zu) ausgewählt wurde. Diese Skala erinnert die angehenden Lehrkräfte an die Notengebung aus dem schulischen Kontext, wobei die sechste Stufe (ungenügend) ausgelassen wurde. Daher steigt der Grad der Zustimmung zum Kriterium mit einem sinkenden Wert der Skala. Damit der mittlere Skalenpunkt 3 nicht als Fluchtkategorie angekreuzt wird, wurde eine zusätzliche legitime Fluchtkategorie (X = weiß ich nicht) dargeboten (nach Porst, 2011). Die Breite der Skala beträgt fünf Skalenpunkte, damit die angehenden Lehrkräfte jeweils rechts und links vom mittleren Skalenpunkt (3) jeweils zwei Ausprägungen haben (1, 2 und 4, 5), in der sie differenziert angeben können, ob das Kriterium nur eher zutrifft (2) oder nicht (4) oder vollständig zutrifft (1) oder nicht (5). Die statistischen Berechnungen, die im Zuge der Hauptuntersuchung durchgeführt wurden, sind in Kapitel 5.5 zu finden.

Während die fünfstufig skalierten Items die Übereinstimmung mit einem Kriterium abfragen, enthält der Bewertungsbogen auch Items, bei denen nach Klassifizierungen gefragt wird (z. B. nach der Repräsentationsform und -ebene). Statistisch gesehen können diese beiden Skalen (fünfstufige Ordinalskala und zweistufige Nominalskala) nicht miteinander verrechnet und in einem statistischen Maß zusammengefasst werden (s. u., nach Porst, 2011). Da die Klassifizierungen in dem Bewertungsbogen ohnehin eine richtungsgebende Funktion haben (z. B. „Kommt in meiner Abbildung die submikroskopische Ebene vor?“)

und zu den Kriterien führen („Wie gut ist die submikroskopische Ebene dargestellt?“), sind sie für das statistische Maß irrelevant und werden bei den Berechnungen ausgelassen. Mit der Untersuchung dieser Auslassung könnten allerdings Forschungsfragen beantwortet werden, die auf das Klassifizierungsvermögen der angehenden Lehrkräfte eingehen. Ein Einblick in das Antwortverhalten der angehenden Lehrkräfte aus qualitativer Perspektive wurde u. a. in der Masterarbeit von Sengutta (2021) gegeben.

Das Arbeiten mit dem Bewertungsbogen erfordert ein bereits ausgeprägtes Maß an fachdidaktischem Wissen, da die angehenden Lehrkräfte im Bewertungsbogen diverse fachdidaktische Grundlagen direkt anwenden und als Kriterien bewerten sollen. Dies beginnt z. B. mit der Übereinstimmung des Lernmaterials mit dem Lernziel, der Wahl eines geeigneten Kontexts, Qualität der Fachsprache und weitet sich auf sowohl kognitionspsychologische (z. B. intendierte Lese- und Denkrichtung) als auch allgemeindidaktische Kriterien (z. B. Aufgabenformulierung, Lernzieltransparenz) aus. Um die Benutzerfreundlichkeit des Fragebogens zu erhöhen und eventuelle Wissenslücken der angehenden Lehrkräfte zu kompensieren, wurde der Bewertungsbogen mit einem inhaltlichen Glossar unterstützt. Potentielle Fachbegriffe, die zu solchen oder ähnlichen Schwierigkeiten führen könnten, wurden im Zuge der Pilotierung in einem Stimulated-Recall-Interview bei den angehenden Lehrkräften erfragt, identifiziert (Tonyali, 2018) und im Glossar ergänzt.

Um zur Benutzerfreundlichkeit weitere Optimierungen durchzuführen, wurde unter jedem Themenblock (z. B. Übereinstimmung mit Lernziel) ein Mittelwert angegeben. Somit konnten die angehenden Lehrkräfte erkennen, in welchen Bereichen ihr designtes Lehr-Lern-Material Stärken und Schwächen aufweist. Dies bildet neben der Skala selbst weitere Feedbackinformationen. Zu berücksichtigen ist jedoch, dass einige Items bzw. Kriterien miteinander zusammenhängen. Das bedeutet, dass wenn ein Kriterium (z. B. fachliche Korrektheit des Texts) nur schwach zutrifft, ein anderes Kriterium (z. B. fachsprachliche Korrektheit des Texts) nur bedingt zutreffen kann. So gibt es einige Unterthemenblöcke wie z. B. die Gliederung und Ordnung des Texts und die sprachliche Angemessenheit oder die Qualität der Aufgabenstellungen und die Lernzieltransparenz.

Um Informationen über die Benutzerfreundlichkeit des Bewertungsbogens zu erhalten, wurde am Ende jeder internen Bewertung der Cognitive Load abgefragt, der während des Arbeitens mit dem Bewertungsbogen von den Proband:innen wahrgenommen wurde. Eine weitere Untersuchungsvariable ist die Anstrengungsbereitschaft, die die angehenden Lehrkräfte einerseits während des Designens der Lehr-Lern-Materialien und andererseits während des Ankreuzens des Bewertungsbogens zeigten. Beide Variablen wurden aus Items aus den Arbeiten von Koenen (2014) entnommen und mit angepasster Formulierung eingesetzt (Kapitel 5.5.2 und 5.5.3).

Das Geben eines standardisierten externen Feedbacks und die Gewährleistung der Objektivität war eine große Herausforderung zumal die Vielfalt an unterschiedlich designten Lehr-Lern-Materialien von der Studienleitung bzw. im externen Feedback gleichermaßen und in maximaler Objektivität bewertet werden sollte. So wurde eine Anleitung für das externe Feedback bzw. ein Kodiermanual erstellt, welches Auskunft darüber gibt, welches Item in welcher Situation mit welcher Ausprägungsform (Skala 1-5) angekreuzt werden soll. Zusätzlich wurde die Transparenz und Objektivität des Kodiermanuals durch die Untersuchung der Interraterreliabilität bzw. Inter-Koder-Übereinstimmung analysiert. Hierzu wurden drei Personen als Kodierer bestimmt: Die Studienleitung bewertete alle in der Intervention designten Lehr-Lern-Materialien, während zwei Hilfskräfte jeweils 20% der Lehr-Lern-Materialien mit Hilfe des Kodiermanuals bewerteten. Es wurde eine Übereinstimmung zwischen den Kodierern mit einem Fleiss' Kappa-Wert von $\kappa = .84$ bestimmt, welche als gute Übereinstimmung bewertet werden kann und somit auf eine hohe Objektivität hinweist (Döring & Bortz, 2023).

Darüber hinaus erfüllt der Bewertungsbogen eine doppelte Funktion innerhalb der Intervention. Zum einen dient er wie bereits beschrieben als Grundlage für das interne und externe Feedback. Zum anderen wird der Bewertungsbogen – und damit das standardisiert generierte externe Feedback – als Instrument zur Erfassung der Qualität der von den angehenden Lehrkräften designten Lehr-Lern-Materialien genutzt und ist somit Teil der Datenerhebung. Die Rolle des Bewertungsbogens als Bewertungshilfe und Testinstrument zugleich ist somit zentral für die Ergebnisse der Studie.

Tabelle 13: Vergleich der Ersatztätigkeiten der Untersuchungsgruppen für das interne und externe Feedback

Untersu- chungsgruppe	Tätigkeit der angehenden Lehrkräfte	
	<i>Internes Feedback</i>	<i>Externes Feedback</i>
Interventions- gruppe 1	Bewertung des selbstdesign- ten Lehr-Lern-Materials mit Hilfe des Bewertungsbogens	Erhalt des von der Studienle- itung ausgefüllten Bewer- tungsbogens für das zugehörige Lehr-Lern-Mate- rial inkl. Teil- und Gesamt- durchschnittsnoten
Interventions- gruppe 2	Bewertung des selbstdesign- ten Lehr-Lern-Materials mit Hilfe des Bewertungsbogens	Erhalt des Dummy-Feed- backs bzw. der Gesamt- durchschnittsnote für das zugehörige Lehr-Lern-Mate- rial
Kontroll- gruppe	Erhalt themenidentischer Schulbuchauszüge und Ver- gleich mit selbstdesigntem Lehr-Lern-Material	Erhalt des Dummy-Feed- backs bzw. der Gesamt- durchschnittsnote für das zugehörige Lehr-Lern-Mate- rial

Anmerkung. Bei den grau hinterlegten Tätigkeiten handelt es sich um Ersatztätigkeiten für systematisches internes oder externes Feedback via Bewertungsbogen.

Zwischen den drei Untersuchungsgruppen gab es Unterschiede bezogen auf die Anwendung des Bewertungsbogens (Tabelle 13): Während die Interventionsgruppe 1 (internes und externes Feedback) den vollständigen Bewertungsbogen nutzte und auch das externe Feedback einsehen konnte, konnte die Kontrollgruppe, welche nur mit dem Dummy-Feedback arbeitete, die Themenabschnitte und einzelnen Kriterien nicht einsehen. Ihnen wurde lediglich der letzte Teil des Bewertungsbogens (Rückmeldung zum Arbeitsblatt) mitgeteilt, wodurch sie abschätzen konnten, in welchem Notenbereich sie sich befinden. Dabei resultierte das Dummy-Feedback durch die (von den Testpersonen nicht einsehbare)

Durchführung des externen Feedbacks und der Zuordnung der Gesamtdurchschnittsnote in eines der vier vorbereiteten Textfragmente (Anhang A.4, „Rückmeldung zum Arbeitsblatt“). Ähnlich wurde auch mit der Interventionsgruppe 2 (internes Feedback) vorgegangen. Die zugehörigen Testpersonen erhielten zwar den Bewertungsbogen und die Kriterien und schätzten sich selbst ein. Sie haben als externes Feedback aber nicht die Bewertung der Studienleitung einsehen können, sondern haben das gleiche Dummy-Feedback wie die Kontrollgruppe erhalten.

4.2.6 Selbsteinschätzungsfähigkeit/Qualität des internen Feedbacks

Der in Kapitel 4.2.5 beschriebene Bewertungsbogen dient nicht nur der Erfassung der Qualität der designten Lehr-Lern-Materialien, sondern auch der Erfassung der Selbsteinschätzungsfähigkeit der angehenden Lehrkräfte. Diese Variable spiegelt wider, wie gut die angehenden Lehrkräfte die Qualität ihrer eigenen Lehr-Lern-Materialien einschätzen können. Um diese Variable zu bestimmen, wurden die Lehr-Lern-Materialien von der Studienleitung mit Hilfe des Kodiermanuals kodiert und als Baseline genutzt (welche gleichzeitig in der Intervention auch den angehenden Lehrkräften als externes Feedback dargeboten wurde). Das interne Feedback der angehenden Lehrkräfte wird hiermit Item für Item verglichen und es wird die Übereinstimmung des internen Feedbacks mit dem externen statistisch ermittelt. Ein hoher Übereinstimmungswert wird dabei mit einer hohen Selbsteinschätzungsfähigkeit bzw. hohen Qualität des internen Feedbacks gleichgesetzt. Da nur die Interventionsgruppen 1 und 2 für das interne Feedback mit dem Bewertungsbogen arbeiteten, liegen die Daten zur Selbsteinschätzungsfähigkeit nur für diese beiden Untersuchungsgruppen vor.

Um das Maß an Übereinstimmung zwischen den Ratern anzugeben, gibt es verschiedene Übereinstimmungskoeffizienten. Dabei gibt ein Koeffizient nicht nur an, ob zwei unterschiedliche Ratings miteinander übereinstimmen, was beispielsweise bei einer prozentualen Übereinstimmung der Fall wäre, sondern auch wie stark die beiden Ratings voneinander abweichen. Beispielsweise wäre bei dem vorliegenden Bewertungsbogen eine Abweichung zwischen 3 und 1 (= -2) größer als zwischen 3 und 2 (= -1). Ebenso zu berücksichtigen ist die Richtung der Abweichung. Eine Abweichung zwischen 3 und 2 (= -1) gibt eine andere Bewertungstendenz an als 3 und 4 (= +1). Um einerseits diese Aspekte zu

berücksichtigen und andererseits Wahrscheinlichkeiten von Zufallsübereinstimmungen mitzuberechnen, gibt es unterschiedliche Übereinstimmungskoeffizienten. Je nach Skalenniveau (Nominal-, Ordinalskala, Intervallskala), Skalenbreite sowie Anzahl an Ratern und Items, über die eine Übereinstimmung berechnet wird, eignet sich ein anderer Übereinstimmungskoeffizient.

In dem vorliegenden Bewertungsbogen wurde eine fünfstufige Likert-Skala eingesetzt. Die Übereinstimmung sollte zwischen dem internen und externen Feedback berechnet werden, sodass zwei Rater vorliegen. In diesem Fall eignet sich beispielsweise das Goodman und Kruskals Gamma (γ) zur Berechnung und Angabe der Übereinstimmung (Naroll, 1974; Somers, 1962). Im Gegensatz dazu wurde beispielsweise für die Analyse der Qualität des Kodiermanuals in Kapitel 4.2.5 das Fleiss' Kappa als Übereinstimmungskoeffizient genutzt, da hier drei Rater einbezogen wurden.

Generell nehmen Reliabilitätskoeffizienten Werte zwischen 1.0 = maximale Übereinstimmung und -1.0 = minimale Übereinstimmung an und können wie in Tabelle 14 klassifiziert werden (Döring & Bortz, 2023):

Tabelle 14: Bewertungsnormen für den Reliabilitätskoeffizienten Goodman und Kruskals Gamma (nach Döring & Bortz, 2023)

Klassifikation	Gut	Mittel	Gering	Unzureichend
Koeffizient	> .80	.60–.80	.40–.60	< .40

Wird der beschriebene Übereinstimmungskoeffizient zur Bestimmung der Selbsteinschätzungsfähigkeit der angehenden Lehrkräfte herangezogen, so indizieren hohe Gamma-Werte eine hohe Selbsteinschätzungsfähigkeit und niedrige Gamma-Werte eine geringe Selbsteinschätzungsfähigkeit.

4.3 Durchgeführte Zusammenhangsanalysen

4.3.1 *Dokumentation der deskriptiven Statistik und Testkennwerte*

Die Beantwortung der in Kapitel 3 beschriebenen Forschungsfragen bedingt eine präzise Dokumentation der Stichprobenmerkmale und deskriptiven Statistik, um diese in der Inferenzstatistik weiterverwenden zu können. Unter den Stichprobenmerkmalen (Tabelle 4) wurden zunächst die demographischen Daten als relative Häufigkeiten wiedergegeben. Für Variablen, die über Likert-skalierte Fragebögen erfasst wurden, wie die persönlichen Ressourcen, Erfahrungen und Überzeugungen, wurden Mittelwerte und Standardabweichungen genutzt. Die Leistungstests, wie das allgemeine Fachwissen, allgemeine fachdidaktische Wissen, repräsentationsbezogene Fachwissen sowie repräsentationsbezogene fachdidaktische Wissen, wurden über relative Summencores berichtet. Dazu wurden die erzielten absoluten Punktwerte pro Test als relativer bzw. prozentualer Anteil der maximal erreichbaren Punkte berechnet. Die Vorgehensweise ermöglicht es, Testergebnisse mit unterschiedlich hoher maximaler Punktzahl vergleichbar darzustellen. Dadurch können zum Beispiel die Ergebnisse des Tests zum allgemeinen Fachwissen (35 Punkte) mit dem zum allgemeinen fachdidaktischen Wissen (42 Punkte) trotz unterschiedlicher maximal möglicher absoluter Summenscores direkt miteinander verglichen werden.

Darüber hinaus wurden zu jedem Fragebogen oder Testinstrument die grundlegenden Testkennwerte angegeben, um u. a. die Aussagekräftigkeit der Daten zu bestimmen (z. B. Tabelle 24). Diese umfassen die interne Konsistenz des Testinstruments und Normalverteilung des Datensatzes. Die Normalverteilung lässt sich visuell durch ein Histogramm abbilden (z. B. Abbildung 18). Liegt im Datensatz eine Normalverteilung vor, so ist diese zunächst optisch über eine symmetrische und eine Glockenkurve bildende Häufigkeitsverteilung erkennbar (Döring & Bortz, 2023). Eine rechnerische Bestätigung der Normalverteilung kann für die vorhandenen Stichprobengrößen durch den Shapiro-Wilk-Test erfolgen (Field, 2024). Zentrale Kennwerte beim Normalverteilungstest sind hierbei die Signifikanz des Shapiro-Wilk-Tests sowie die z-standardisierte Schiefe und Kurtosis. Diese geben Auskunft über die Form der Verteilung und sollten Werte um 0 betragen. Liegt die Signifikanz des Tests bei $p > .05$, kann die Nullhypothese, dass die Daten normal verteilt sind, nicht abgelehnt werden. Dieses Ergebnis wäre ein Indiz dafür, dass die Daten einer Normalverteilung entsprechen

und die Ergebnisse als repräsentativ für die Grundgesamtheit der Stichprobe herangezogen werden können (Field, 2024; Shapiro & Wilk, 1965).

4.3.2 *Varianzanalysen*

In der vorliegenden Studie werden Varianzanalysen benötigt, um potenzielle Gruppenvarianzen zwischen den drei Untersuchungsgruppen zu untersuchen. So werden mit allen Kontrollvariablen und abhängigen Variablen (Tabelle 4) Varianzanalysen durchgeführt, um einerseits die Vergleichbarkeit der Untersuchungsgruppen zu gewährleisten und andererseits Effekte der Intervention zu analysieren. Auch für Varianzanalysen gibt es eine Auswahl an unterschiedlichen Tests, die von der Normalverteilung und Art des Vergleichs abhängt. Wird beispielsweise die Leistung einer einzigen Stichprobe (z. B. Interventionsgruppe 1) zu zwei verschiedenen Messzeitpunkten (z. B. Prä- und Posttest) miteinander verglichen, so eignet sich eine Varianzanalyse für verbundene Stichproben. Wird aber die Leistung zweier unterschiedlicher Gruppen (z. B. Interventionsgruppe 1 und 2) zum selben Messzeitpunkt (z. B. Posttest) miteinander verglichen, so sollten Varianzanalysen für unabhängige Stichproben herangeführt werden (Field, 2024; Schwarz, 2023). Das bedeutet, dass zur Beantwortung der Forschungsfragen mehrere Arten von Varianzanalysen eingesetzt wurden, welche in der Abbildung 7 als Übersicht zusammengetragen sind. In den Beispielergebnissen der Abbildung 7 wie auch in der Dokumentation der Ergebnisse (Kapitel 5) sind je nach eingesetztem Test unterschiedliche Parameter zu erkennen: Der erste Teil (z. B. U , Z , χ^2 , t , F) bezeichnet die Stärke der Varianz, wobei der p -Wert die Signifikanz dieses Ergebnisses angibt. Ein p -Wert kleiner als .05 deutet auf ein signifikantes Ergebnis hin, was bedeutet, dass das Ergebnis mit weniger als 5 % Wahrscheinlichkeit zufällig ist. Auf diese Weise kann beispielsweise überprüft werden, ob eine Untersuchungsgruppe signifikant höhere Leistungen erbracht hat als eine andere. Die Effektstärke, die die Größe des Unterschieds oder Effekts angibt, wird durch einen weiteren Zahlenwert dargestellt. Die Effektstärke kann je nach Art der durchgeführten Analyse in unterschiedlichen Maßeinheiten angegeben werden (Tabelle 15). Die unterschiedlichen Maße für die Effektstärken weisen unterschiedliche Bewertungsnormen auf. Die Maße können jedoch in andere Effektstärken umgerechnet werden, was hilfreich sein kann, wenn die Effektstärken miteinander verglichen werden sollen (Lenhard & Lenhard, 2017). Eine vor allem bei ANOVA häufig

verwendete Kennzahl ist das Eta-Quadrat (η^2), welches in diesem Fall ausgewählt wurde, um im Ergebnisteil alle Effektstärken einheitlich wiederzugeben (nach Field, 2024).

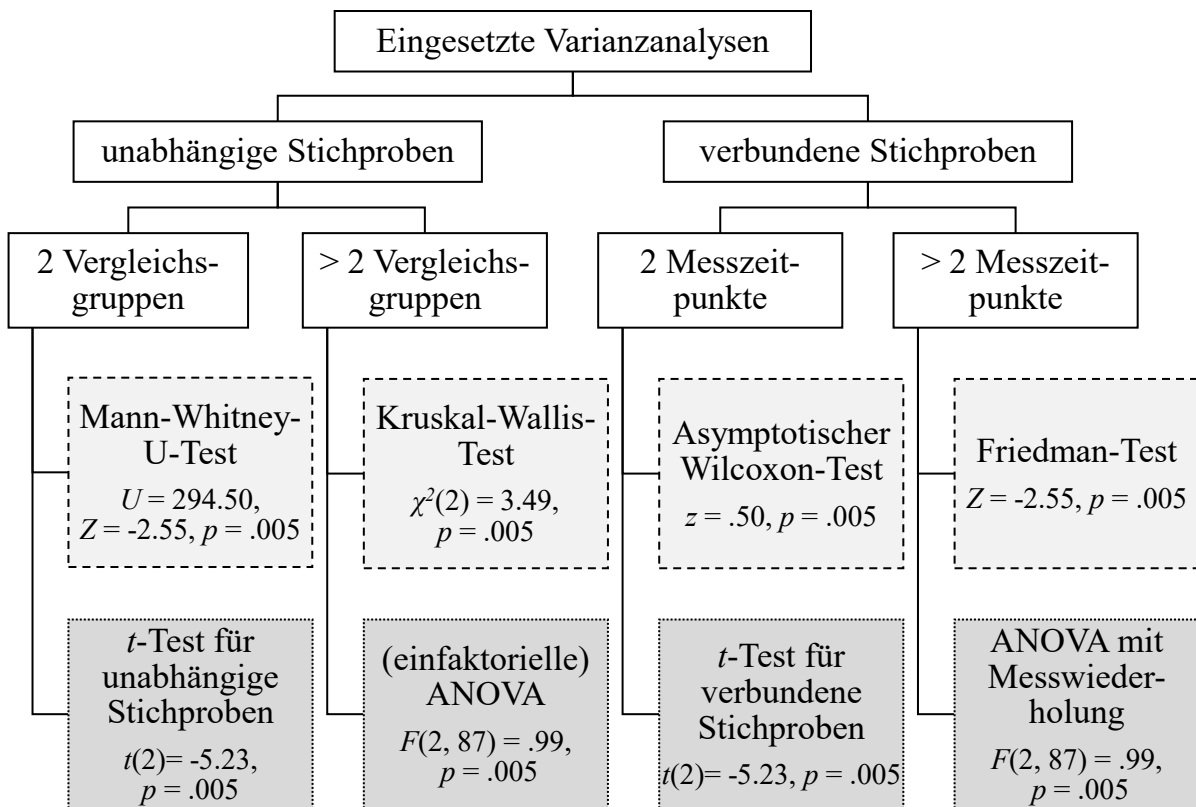


Abbildung 7: Entscheidungspfad zur Auswahl der passenden Varianzanalysen für den vorliegenden Datensatz mit Beispielergebnissen

Anmerkung. Die punktiert eingerahmten Tests eignen sich für normalverteilte Daten. Die gestrichelt eingerahmten Tests eignen sich für nicht normalverteilte Daten (erstellt nach Field, 2024; Schwarz, 2023). Die berechneten ANOVA wurden mit Bonferroni-korrigierten post-hoc Tests ergänzt (z. B.: MDiff=9.61, 95%-CI[.50, 18.71], $p = .035$; $t(59) = 2.48$, $p = .035$, $\eta^2 = .10$), um die genaue Gruppenvarianz anzugeben (Field, 2024). Die ANOVA mit Messwiederholung wurden mit Greenhouse-Geisser-Korrektur berechnet (Field, 2024; Schwarz, 2023).

Tabelle 15: Übersicht zur Einordnung von Effektstärken bei Varianzanalysen und Korrelationen

Art der Analyse	Maß der Effektstärke	Formel	Effekt		
			klein	mittel	groß
t-Test	Cohens d	$d = \frac{2t}{\sqrt{df}}$	< .5	.5-.8	> .8
ANOVA	Eta Quadrat η^2	$\eta^2 = \frac{QS_{Zwischen}}{QS_{Gesamt}}$	< .06	.06-.14	> .14
Mann-Whitney Test Wilcoxon Test Kruskal-Wallis Test Pearson Korrelation Spearman Korrelation	Pearson Korrelationskoeffizient r	$r = \frac{z}{\sqrt{N}}$	< .3	.3-.5	> .5
Chi-Quadrat Test	Phi Φ	$\phi = \sqrt{\frac{\chi^2}{N}}$	< .25	.25-.66	> .66
	Cramers V	$V = \sqrt{\frac{\chi^2}{N \cdot m}}$	< .3	.3-.4	> .4

Anmerkung. df = Freiheitsgrad, m = Minimum von N und χ , N = Stichprobengröße, QS_{Gesamt} = Fehlerquadratsumme Gesamt, $QS_{Zwischen}$ = Fehlerquadratsumme zwischen den Gruppen, t = t -Wert, z = standardisierter z -Wert. Erstellt nach Field, 2024; Lenhard & Lenhard, 2017.

Bei den nominalskalierten Variablen (z. B. Häufigkeit des Geschlechts), wurden ebenso Varianzanalysen durchgeführt, um beispielsweise bei den demographischen Daten auszuschließen, dass bestimmte Stichprobeneigenschaften in einer Untersuchungsgruppe besonders stark vertreten sind. Da hier keine Normalverteilung bestimmt werden kann, wurden Chi-Quadrat-Tests durchgeführt. (z. B.: $\chi^2(6) = 2.53, p = .865$; Field, 2024).

4.3.3 Bivariate Zusammenhangsanalysen (Korrelation)

Eine weitere Analyse, um Zusammenhänge zwischen den erfassten Untersuchungsvariablen aufzudecken, ist die Korrelationsanalyse. Korrelationsanalysen wurden in der vorliegenden Arbeit eingesetzt, um beispielsweise Zusammenhänge zwischen den persönlichen Ressourcen und den Leistungen in den Wissensarten zu ermitteln. Eine Korrelation beschreibt den Zusammenhang zwischen zwei oder mehr Variablen und gibt an, in welchem Maße und in welcher Richtung diese miteinander in Beziehung stehen. Der Wert des Korrelationskoeffizienten liegt im Bereich von -1 bis +1, wobei ein Wert von +1 eine perfekte positive Korrelation bedeutet. Dies deutet darauf hin, dass ein Anstieg der einen Variablen stets mit einem gleichmäßigen Anstieg der anderen Variablen einhergeht. Ein Wert von -1 zeigt eine perfekte negative Korrelation an, was bedeutet, dass ein Anstieg der einen Variablen mit einem gleichmäßigen Abfall der anderen Variablen verbunden ist (Döring & Bortz, 2023). Die Stärke der Korrelation wird durch den Betrag des Korrelationskoeffizienten bestimmt (Tabelle 16; Field, 2024).

Tabelle 16: Zur Auswertung ausgewählte Bewertungsnormen für den Reliabilitätskoeffizienten Cronbachs Alpha

Klassifikation	Sehr schwach	Schwach	Mittel	Stark	Sehr stark
$ r $	.00–.10	.10–.30	.30–.50	.50–.70	.70–1.00

Es ist jedoch wichtig zu beachten, dass eine Korrelation nicht gleichbedeutend mit Kausalität ist. Auch wenn zwei Variablen stark korrelieren, bedeutet dies nicht direkt, dass die eine Variable eine Veränderung der anderen verursacht. Eine hohe Korrelation kann durch einen Einfluss von Drittvariablen oder Mediatoren entstehen oder durch andere nicht erfasste Faktoren bedingt sein. Darüber hinaus erfasst der Korrelationskoeffizient nur lineare Beziehungen zwischen den Variablen und ist nicht geeignet, um nicht-lineare Zusammenhänge zu identifizieren. Streudiagramme können hier nützlich sein, um diese aufzudecken (Kapitel 5.6.3). Somit stellt die Korrelation ein nützliches, aber auch begrenztes Maß dar, um Beziehungen zwischen Variablen zu verstehen und sollte immer

im Kontext der Daten und der Forschungsfrage interpretiert werden (Döring & Bortz, 2023).

4.3.4 Multivariate Zusammenhangsanalysen (Multiple lineare Regression)

Die multiple lineare Regression ist ein statistisches Verfahren, welches ebenso wie die Korrelationsanalyse Zusammenhänge untersucht, aber noch weitere Informationen zur Interpretation bietet. Im Gegensatz zur einfachen Korrelation, die lediglich die Stärke und Richtung des linearen Zusammenhangs zwischen zwei Variablen misst, erlaubt die Regression eine detailliertere Analyse, indem sie nicht nur den Zusammenhang zwischen den Variablen betrachtet, sondern auch die Möglichkeit bietet, Vorhersagen zu treffen (Döring & Bortz, 2023).

In der multiplen linearen Regression können mehrere unabhängige Variablen gleichzeitig berücksichtigt werden, wodurch eine Modellierung und Analyse komplexer Beziehungen ermöglicht wird. Bei der Berechnung wird eine abhängige Variable ausgewählt (z. B. das repräsentationsbezogene Fachwissen). Anschließend werden eine oder mehrere unabhängige Variablen ausgewählt, die möglicherweise einen Einfluss auf die unabhängige Variable haben können (z. B. das allgemeine Fachwissen, das repräsentationsbezogene fachdidaktische Wissen, der wahrgenommene Zeitdruck, die Selbstwirksamkeitserwartung). Bisher festgestellte Korrelationen (Kapitel 5.6.2) können dabei indizieren, welche unabhängigen Variablen dabei in Frage kommen. Danach werden die ausgewählten unabhängigen Variablen in das Regressionsmodell einbezogen, sodass ihr Einfluss als Prädiktoren über errechnete Parameter interpretiert werden kann (IBM Corp., 2016). Dabei kann so vorgegangen werden, dass erst der stärkste Prädiktor einbezogen und das Modell nach und nach in mehreren Stufen mit den nächst stärksten Prädiktoren ergänzt wird (Methode: Stufenweise, z. B. Tabelle 63). Wie stark ein Prädiktor ist, gibt der zugehörige standardisierte Beta-Koeffizient (β) an (Tabelle 17). Würde zum Beispiel der standardisierte Beta-Koeffizient des allgemeinen Fachwissens $\beta = 0.80$ betragen, so würde das bedeuten, dass sich das repräsentationsbezogene Fachwissen im Mittel um 0.80 Standardabweichungen erhöht, wenn das allgemeine Fachwissen um eine Standardabweichung zunimmt. Zusammen mit dem Beta-Koeffizienten wird auch der zugehörige p -Wert berichtet, aus dem die Signifikanz des Prädiktors interpretiert wird (Field, 2024).

Tabelle 17: Bewertungsnorm für R^2 zur Bewertung der Varianzaufklärung eines Regressionsmodells

Effektgröße	Sehr schwach	Schwach	Moderat	Stark
β	< .10	.10–.30	.30–.50	> .50

Darüber hinaus werden das passende Regressionsmodell und die darin enthaltenen unabhängigen Variablen in der Regel nach einer möglichst hohen Modellgüte ausgewählt. Beispielsweise wird hierfür der multiple Determinationskoeffizient (R^2) genutzt (Tabelle 18; Cohen, 1988). Weitere Eigenschaften des Modells werden ähnlich wie bei den Varianzanalysen angegeben (z. B.: $F(82, 3) = 8.57, p < .001, N = 90$; Field, 2024). Um das ermittelte Regressionsmodell zu berichten und zu visualisieren eignen sich vor allem Pfaddiagramme, wie zum Beispiel in der Abbildung 40.

Tabelle 18: Bewertungsnorm für R^2 zur Bewertung der Varianzaufklärung eines Regressionsmodells

Varianzaufklärung	Schwach	Moderat	Stark
$ R^2 $	$\leq .02$	.13	$\geq .26$

5 ERGEBNISSE DER HAUPTSTUDIE

5.1 Stichprobenmerkmale der angehenden Lehrkräfte

5.1.1 *Demografische Daten*

An der Hauptstudie nahmen 90 angehende Lehrkräfte aus 19 ZfsL bzw. 21 Chemiefachseminaren teil. Die Anzahl der Teilnehmenden aus den einzelnen Seminaren unterschied sich, weil die Seminargruppen unterschiedlich groß waren. Von insgesamt 50 infrage kommenden Chemiefachseminaren in NRW der Lehramtsformen HRSGe und GymGe (Stand: Februar 2020) meldeten sich 20 Seminare für die Teilnahme an der Studie an. Aus den Anmeldeunterlagen der zugehörigen Fachseminarleitungen ging hervor, dass der Studie schätzungsweise 210 angehende Lehrkräfte als potenzielle Testpersonen zur Verfügung standen. Zur Teilnahme an der Studie meldeten sich von dieser Gesamtheit 122 Testpersonen an, von denen allerdings 26 % den Prä- und Posttest sowie die Intervention nicht oder nicht vollständig absolvierten und daher von der Auswertung ausgeschlossen wurden. Letztlich lagen 90 vollständige und bereinigte Datensätze von angehenden Chemielehrkräften als Datengrundlage vor.

Die Testpersonen wurden möglichst gleichmäßig auf die drei Untersuchungsgruppen verteilt. Bei der Zuordnung der Testpersonen zu einer Untersuchungsgruppe standen nur erste Informationen aus den Anmeldeunterlagen zur Verfügung, welche den Standort des Chemiefachseminars, die Lehramts- und Schulform und den Dienstgrad (Lehramtsanwärter:in oder Seiteneinsteiger:in) beinhalteten. Diese Informationen und weitere demografische Merkmale sind jeweils für die Gesamtgruppe und die drei Untersuchungsgruppen in Tabelle 19 aufgeführt. Die weiteren demografischen Merkmale wurden erst nach der Gruppenzuordnung im Prätest erfragt, sodass die einzelnen Gruppenmerkmale wie Alter und Abschlussnoten randomisiert entstanden sind. Dennoch wurden alle demografischen Daten auf gruppenspezifische Varianzen untersucht bzw. darauf, ob einige Merkmale übermäßig oft in einer Untersuchungsgruppe auftreten und diese in Bezug auf die Forschungsergebnisse eine Rolle spielen. Hierzu wurde je nach Datenstruktur ein Chi-Quadrat-Test, ein Kruskal-Wallis-Test oder eine ANOVA durchgeführt. Bei allen Stichprobenmerkmalen wurden keine gruppenspezifischen Varianzen festgestellt (Anhang C, Tabelle 57), sodass davon ausgegangen werden kann, dass sich alle abgebildeten Variablen homogen

verteilen. Es ist daher möglich, Varianzanalysen der unabhängigen Variablen zwischen den Untersuchungsgruppen durchzuführen.

Des Weiteren wurden Details zum Hochschulstudiengang erfasst, mit denen sich die angehenden Lehrkräfte für den Vorbereitungsdienst qualifiziert haben. Hierbei handelt es sich um die Form des Studiengangs und das Bundesland, in dem der Hochschulabschluss erworben wurde. Aus dem Kreisdiagramm in Abbildung 8 wird ersichtlich, dass die Stichprobe zu etwa zwei Dritteln aus grundständigen Lehramtsstudierten und einem Drittel aus Seiten- und Quereinsteiger:innen besteht. Außerdem haben 95 % der Gesamtgruppe den Hochschulabschluss in NRW absolviert, während 5 % diesen außerhalb von NRW erhalten haben.

Tabelle 19: Stichprobenzusammensetzung

Stichprobenmerkmal		Untersuchungsgruppe			
Merkmal	Kategorien	IG1	IG2	KG	Insgesamt
<i>N</i> (Personen)		29	29	32	90
<i>N</i> (ZfsL)		16	17	15	18
Lehramtsform [%]	HRSGe	41	45	31	39
	GymGe	59	55	69	61
Schulform [%]	Hauptschule	0	7	6	4
	Realschule	7	4	13	8
	Sekundarschule	3	7	6	6
	Gesamtschule	49	41	19	35
	Gymnasium	41	41	56	47
Dienstgrad [%]	LAA	62	72	66	67
	LiA	38	28	34	33
Absolvierte Zeit im Vorbereitungsdienst [%]	1-3 Monate	31	32	34	32
	4-15 Monate	41	54	60	52
	16-18 Monate	14	7	3	8
	19-24 Monate	14	7	3	8

Dauer einer vorherigen Vertretungsstelle [%]	0 Monate	52	69	59	60
	3 Monate	10	7	9	9
	12 Monate	14	14	13	13
	≥ 24 Monate	24	10	19	18
Alter [%]	21-25 Jahre	14	14	12.5	13
	26-30 Jahre	48	45	53	49
	31-35 Jahre	17	17	12.5	16
	36-40 Jahre	14	14	19	16
	≥ 41 Jahre	7	10	3	6
Abschlussnote im Abitur	[M]	2.12	2.12	2.17	2.14
	[SD]	.59	.61	0.59	.59
Hochschul- abschlussnote	[M]	1.71	1.91	1.84	1.82
	[SD]	.38	.64	.43	.50
Muttersprache	Deutsch	90	93	75	85
	Deutsch u. An- dere	7	3.5	6	6
	Andere	3	3.5	19	9
	weiblich	62	62	69	64
Geschlecht	männlich	38	38	31	36
	divers	0	0	0	0

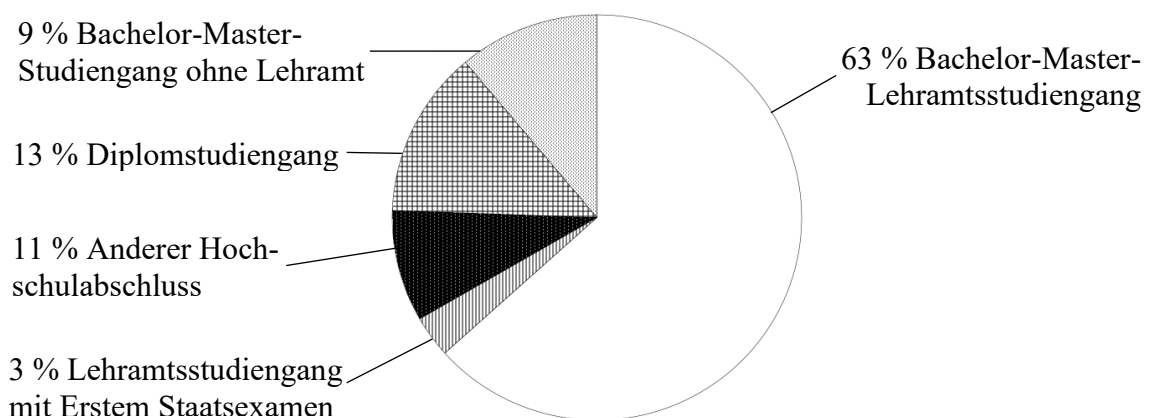


Abbildung 8: Absolvierte Studiengänge der angehenden Lehrkräfte

Verglichen mit den Ergebnissen der Skalen zur Qualität der Lehrkräftebildung von Abs et al. (2005) sind die Stichprobenmerkmale nahezu identisch. Dies gilt insbesondere für Alter, Geschlecht, Schulform und Phase des Vorbereitungsdienstes.

5.1.2 Persönliche Ressourcen im Vorbereitungsdienst

Als weitere Stichprobenmerkmale wurden die persönlichen Ressourcen der angehenden Lehrkräfte untersucht. Die Kennwerte der Subskalen indizieren eine Normalverteilung sowie starke bis sehr starke Reliabilitäten (nach Field, 2024). Außerdem kommen die vorliegenden Werte den zugehörigen Reliabilitätswerten der originalen Skalen von Abs et al. (2005) sehr nahe, welche bei $\alpha = .61$ und $\alpha = .86$ liegen.

Tabelle 20: Testkennwerte zu den persönlichen Ressourcen

Variable	<i>n</i> (Items)	Shapiro- Wilk-Test [<i>p</i>]	Schiefe (<i>z</i> -stand.)	Kurtosis (<i>z</i> -stand.)	Reliabilität [α]
<i>Belastungen und Bewältigungsstrategien</i>					
Zeitdruck	3	> .050	.02	-.23	.71
Berufli. Belastung	12	> .050	.19	.14	.78
Lernstrategien	7	> .050	-.14	-.44	.60
Elaborationsmuster	5	> .050	-.29	.20	.74
Soziale Ressourcen	8	> .050	-.64	.76	.76
<i>Professionelles Selbstverständnis</i>					
Wichtigkeit der Wissenschaft	6	> .050	-.15	.45	.61
SWE an sich selbst	15	> .050	-.06	-.43	.76
SWE an eigenen Chemieunterricht	15	> .050	.40	.28	.76

Die Daten zu den *Belastungen und Bewältigungsstrategien* wurden durch eine vierstufige Likert-Skala generiert (0 = „keine Ausprägung“ bis 3 = „sehr starke Ausprägung“). Die Subskalen weisen insgesamt eine neutrale bis starke Ausprägung auf (Tabelle 21). Die angehenden Lehrkräfte gaben an, dass sie eine mittelmäßige Belastung erfahren, jedoch gleichzeitig Bewältigungsstrategien in einem mittleren bis starken Maß anwenden. Varianzanalysen zeigen, dass das Belastungsempfinden und das Ausmaß der angewandten Bewältigungsmuster in allen drei Untersuchungsgruppen gleich stark ausgeprägt sind (Zeitdruck: $\chi^2(2) = .64, p = .728$; Berufliche Belastung: $\chi^2(2) = .54, p = .763$; Lernstrategien: $\chi^2(2) = 1.05, p = .592$; Elaborationsmuster: $\chi^2(2) = 4.02, p = .134$; Soziale Ressourcen: $\chi^2(2) = 1.00, p = .995$).

Tabelle 21: Vergleich der mittleren Belastungen und Bewältigungsstrategien

Variable	Untersuchungsgruppe [<i>M, SD</i>]			
	IG1	IG2	KG	Insgesamt
Zeitdruck	1.68, .55	1.60, .64	1.72, .45	1.67, .55
Berufliche Belastung	1.22, .48	1.20, .58	1.20, .43	1.21, .49
Lernstrategien	2.25, .47	2.12, .49	2.15, .52	2.17, .49
Elaborationsmuster	1.82, .45	1.59, .52	1.59, .49	1.66, .49
Soziale Ressourcen	2.11, .51	2.15, .39	2.10, .52	2.12, .47

Anmerkung. Skala: 0 = „keine Ausprägung“ bis 3 = „sehr starke Ausprägung“.

Tabelle 22: Vergleich des mittleren professionellen Selbstverständnisses

Variable	Untersuchungsgruppe [<i>M, SD</i>]			
	IG1	IG2	KG	Insgesamt
Wichtigkeit der Wissenschaft	1.78, .45	2.08, .40	1.89, .35	1.91, .41

Anmerkung. Skala: 0 = „überhaupt keine Wichtigkeit“ bis 3 = „sehr hohe Wichtigkeit“.

Tabelle 23: Vergleich der mittleren individuellen beruflichen Selbstwirksamkeitserwartung

Variable	Untersuchungsgruppe [<i>M</i> , <i>SD</i>]			
	IG1	IG2	KG	Insgesamt
SWE an sich selbst	2.40, .44	2.51, .56	2.59, .51	2.50, .51
SWE an eigenen Chemieunterricht	2.23, .55	2.22, .66	2.25, .51	2.24, .57

Anmerkung. Skala: 0 = „sehr geringe Ausprägung“ bis 4 = „sehr starke Ausprägung“.

Die Subskalen zur Einschätzung der Wichtigkeit der Wissenschaft für die Praxis sowie zur Selbstwirksamkeitserwartung bilden das *professionelle Selbstverständnis* ab. Da die Wichtigkeit der Wissenschaft mit einer vierstufigen (0 = „überhaupt keine Wichtigkeit“ bis 3 = „sehr hohe Wichtigkeit“), die Selbstwirksamkeitserwartung jedoch mit einer fünfstufigen Likert-Skala (0 = „sehr gering“ bis 4 = „sehr stark“) erfasst wurde, werden deren Ergebnisse getrennt berichtet (Tabelle 22, Tabelle 23). Auch bei der Wichtigkeit der Wissenschaft zeigt sich im Mittelwert eine positive Tendenz, die in allen Gruppen gleich stark ausgeprägt ist ($\chi^2(2) = 5.92$, $p = .082$). Verglichen mit den Angaben der angehenden Lehrkräfte in den originalen Skalen von Abs et al. (2005) sind auch die berechneten Mittelwerte nahezu identisch.

Die individuelle berufliche Selbstwirksamkeitserwartung umfasst zwei Subskalen: 1. die Selbstwirksamkeitserwartung an sich selbst (z. B. „Ich bin kompetent genug, um eine gute Lehrkraft zu sein und werde mich immer mehr weiterentwickeln.“) und 2. die Selbstwirksamkeitserwartung an den eigenen Chemieunterricht (z. B. „Wenn ich mich bei der Unterrichtsvorbereitung anstrengte, werden meine Schüler:innen bessere Leistungen erbringen.“). Bei beiden Subskalen zeigen sich neutrale bis leicht positive Erwartungen. Zwischen den Untersuchungsgruppen gibt es keine Varianzen (SWE an sich selbst: $\chi^2(2) = 1.07$, $p = .584$; SWE an eigenen Chemieunterricht: $\chi^2(2) = .01$, $p = .998$). Werden die gemessenen Kenn- und Mittelwerte der Selbstwirksamkeitserwartungen mit der originalen Skala von Bleicher (2004) verglichen, so stimmen die Ergebnisse beider Erhebungen auch hier weitgehend überein. Darüber hinaus wurden hohe Selbstwirksamkeitserwartungen angehender Lehrkräfte auch von Mutke (2017) berichtet. Diese erhöhen sich im Laufe des Vorbereitungsdiensts.

5.1.3 *Erfahrungen in der Gestaltung von Lehr-Lern-Materialien*

Die Erfahrung der angehenden Lehrkräfte in der Gestaltung von Lehr-Lern-Materialien wird in fünf Unteraspekte unterteilt: (1) Anknüpfen an das Vorwissen der Schüler:innen, (2) Vorbereiten von Unterrichtsmaterial, (3) Logisches Gliedern von ausgewählten Lerninhalten, (4) Strukturieren von chemischen Inhalten mit Hilfe des Lehrplans über mehrere Klassenstufen, (5) Auswählen von Unterrichtsmaterial auf Basis eigener Bewertungen. Im ersten Schritt wurde hierzu erhoben, wo die angehenden Lehrkräfte diese Erfahrungen im Vorbereitungsdienst gesammelt haben (Abbildung 9). In den Selbstauskünften wird deutlich, dass sich die angehenden Lehrkräfte bisher vor allem in der Schule mit dem Gestalten von Lehr-Lern-Material auseinandergesetzt haben, beispielsweise beim Vorbereiten des eigenen Unterrichts.

Im zweiten Schritt wurde die Kompetenz der angehenden Lehrkräfte zu den einzelnen Unteraspekten per Selbsteinschätzung anhand einer fünfstufigen Likert-Skala abgefragt (Tabelle 24). Die Mittelwerte zu den Unteraspekten werden als zusammengefasste Variable „Gesamterfahrung“ sowie einzeln berichtet (Tabelle 25).

Insgesamt schätzen die angehenden Lehrkräfte ihre Erfahrungen in einem ausreichenden bis zufriedenstellenden Maß ein. Bezüglich der Unteraspekte sind zwischen den Untersuchungsgruppen keine Varianzen festzustellen (Anknüpfen: $\chi^2(2) = 1.37, p = .504$; Vorbereiten: $\chi^2(2) = 1.96, p = .375$; Gliedern: $\chi^2(2) = .74, p = .689$; Strukturieren $\chi^2(2) = .27, p = .874$; Auswählen $\chi^2(2) = 3.93, p = .140$). Dies gilt auch für die Gesamterfahrung ($F(2, 87) = .996, p = .373$). In den originalen Skalen von Abs et al. (2005) liegen hier keine Kennwerte und Ergebnisse vor, sodass hier kein Vergleich durchgeführt werden kann.

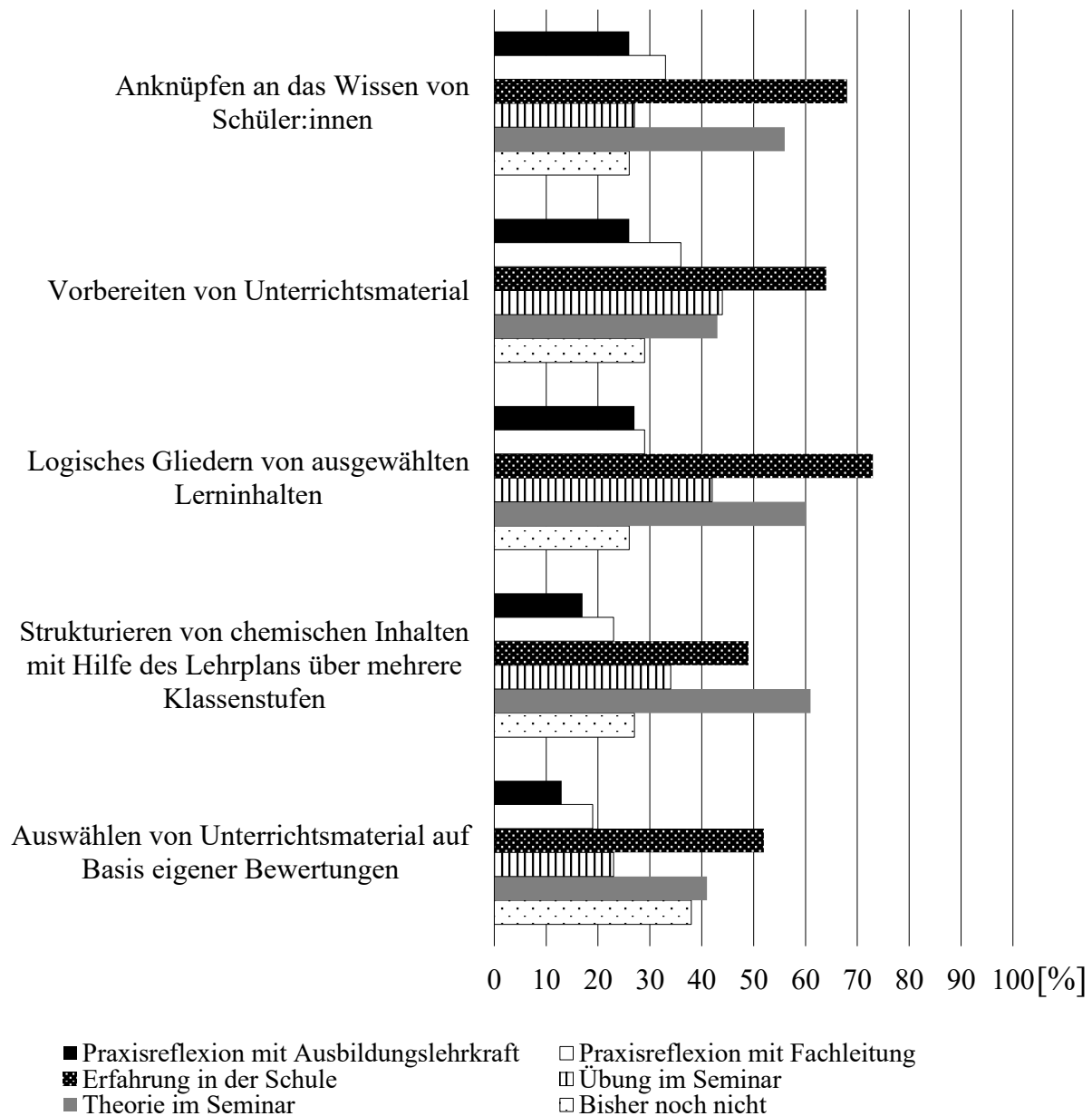


Abbildung 9: Relative Häufigkeiten zum Lernkontext der Gestaltung von Lehr-Lern-Materialien

Tabelle 24: Testkennwerte zu den Erfahrungen in der Gestaltung von Lehr-Lern-Materialien

Variable	Shapiro-Wilk-Test [p]	Schiefe (z-stand.)	Kurtosis (z-stand.)	Reliabilität [α]
Gesamterfahrung	> .050	-.12	-.22	.85

Tabelle 25: Vergleich der mittleren Erfahrungen in der Gestaltung von Lehr-Lern-Material

Variable	Untersuchungsgruppe [<i>M</i> , <i>SD</i>]			
	IG1	IG2	KG	Insgesamt
Anknüpfen an das Vorwissen der Schüler:innen	1.97, .94	1.72, 1.00	2.00, .84	1.90, .92
Vorbereiten von Unterrichtsmaterial	2.31, .76	2.03, .87	2.19, .64	2.18, .76
Logisches Gliedern von ausgewählten Lerninhalten	2.34, .81	2.45, .74	2.34, .70	2.38, .74
Strukturieren von chem. Inhalten mit Hilfe des Lehrplans über mehrere Klassen	1.93, 1.07	1.72, 1.07	1.91, .86	1.86, .99
Auswählen von Unterrichtsmaterial auf Basis eigener Bewertungen	1.97, .94	1.72, 1.00	2.00, .84	1.90, .99
Gesamterfahrung	2.10, .74	1.93, .78	2.09, .61	2.04, .71

Anmerkung. Skala: 0 = „nicht ausreichende Beherrschung“ bis 4 = „ausgezeichnete Beherrschung“.

5.1.4 Allgemeines Fachwissen und fachdidaktisches Wissen

Das allgemeine Fachwissen und fachdidaktische Wissen wurden als Kontrollvariablen im Prä- und Posttest erfasst und werden im Folgenden als mittlere relative Summenscores berichtet. Während die Ergebnisse des Fachwissenstests normalverteilt sind, sind die des fachdidaktischen Wissens es nicht (Tabelle 26). Die zugehörigen Daten zeigen starke bis sehr starke Reliabilitäten.

Tabelle 26: Testkennwerte zum allgemeinen Fachwissen und fachdidaktischen Wissen

Variable	Shapiro-Wilk- Test [p]	Schiefe (z -stand.)	Kurtosis (z -stand.)	Reliabilität [α]
Allgemeines Fachwissen (prä)	> .050	-.14	-.20	.61
Allgemeines Fachwissen (post)	> .050	-.11	-.75	.61
Allgemeines fachdidakti- sches Wissen (prä)	< .001	-1.31	2.20	.87
Allgemeines fachdidakti- sches Wissen (post)	< .001	-1.24	2.63	.82

Tabelle 27: Vergleich des mittleren relativen allgemeinen Fachwissens und fachdidaktischen Wissens

Variable	Untersuchungsgruppe [M , SD]			
	IG1	IG2	KG	Insgesamt
Allgemeines Fachwissen (prä)	58.62, 20.73	59.54, 19.76	55.62, 15.30	57.85, 18.50
Allgemeines Fachwissen (post)	62.30, 17.39	62.07, 14.92	60.63, 20.35	61.63, 17.60
Allgemeines fach- did. Wissen (prä)	55.92, 12.95	60.88, 12.74	59.66, 7.84	58.85, 11.38
Allgemeines fach- did. Wissen (post)	55.31, 12.33	56.76, 13.96	56.81, 9.06	56.31, 11.76

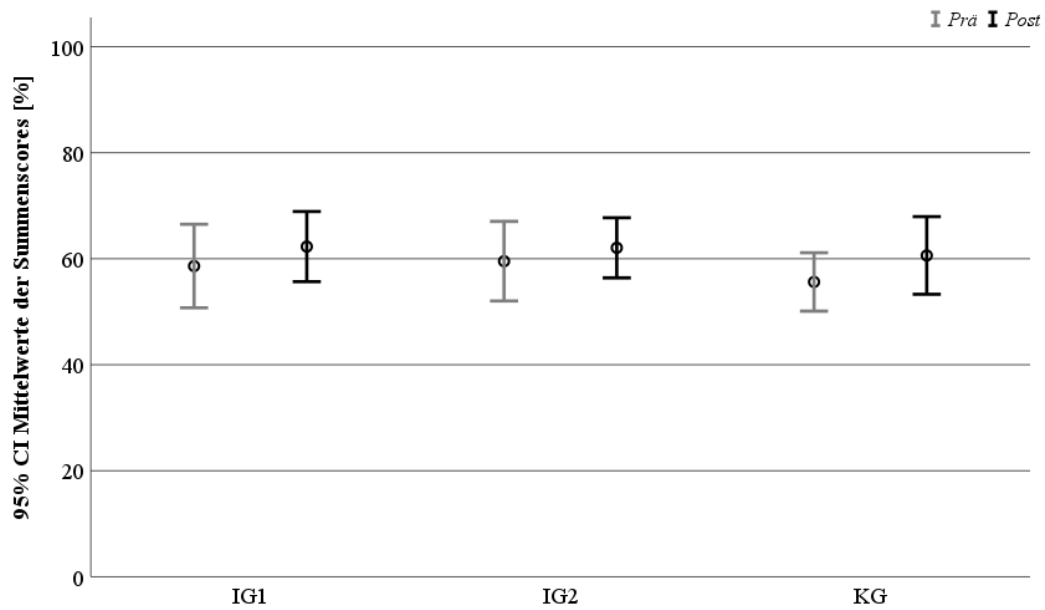


Abbildung 10: Vergleich des mittleren relativen allgemeinen Fachwissens

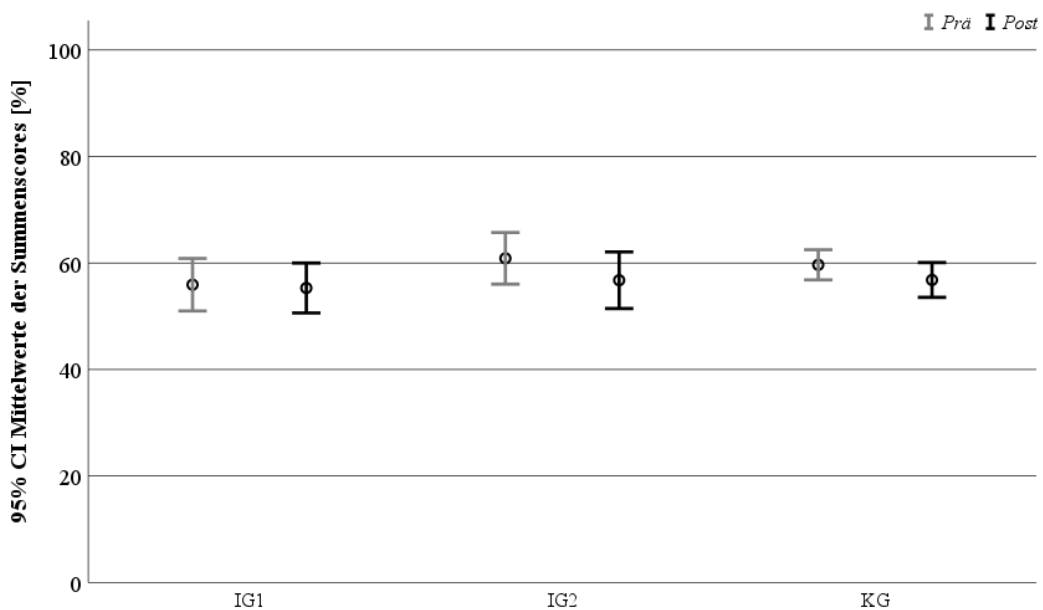


Abbildung 11: Vergleich des mittleren relativen allgemeinen fachdidaktischen Wissens

Die Mittelwerte und zugehörigen Standardabweichungen der beiden Tests sind in Tabellen- und Diagrammform für beide Messzeitpunkte abgebildet (Tabelle 27, Abbildung 10, Abbildung 11). Die Ergebnisse beider Tests liegen insgesamt

in einem ähnlichen Leistungsbereich von 55-60 %. Varianzanalysen zeigen, dass es im Prätest keine Unterschiede zwischen den Untersuchungsgruppen gibt. Dies gilt sowohl für das Fachwissen ($F(2, 86) = 1.17, p = .314$) als auch das fachdidaktische Wissen ($\chi^2(2) = 3.49, p = .175$). Diese Varianzhomogenität bleibt auch im Posttest bestehen (Fachwissen: $F(2, 86) = .08, p = .923$; fachdidaktisches Wissen: $\chi^2(2) = .72, p = .698$).

Bei den Varianzanalysen mit Messwiederholung werden die beiden Wissensbereiche im Prä- und Posttest jeweils miteinander verglichen. Signifikante Veränderungen der Mittelwerte sind im Verlauf der Intervention nicht zu finden (Fachwissen: $F(2, 87) = .19, p = .827$; fachdidaktisches Wissen: IG1 $z = .50, p = .620$, IG2 $z = 2.20, p = .028$, KG $z = 1.89, p = .058$).

Die Stichprobenmerkmale, zusammengesetzt aus den demografischen Daten, persönlichen Ressourcen sowie dem allgemeinen Fachwissen und fachdidaktischen Wissen, wurden berichtet, um zum einen die Merkmale und Eigenschaften der Stichprobe zu ermitteln und mögliche Erklärungen für spätere Ergebnisse herzuleiten. Zum anderen sollte die Vergleichbarkeit der drei Untersuchungsgruppen überprüft werden, da die Testpersonen in Bezug auf ihr Vorwissen und Erfahrungen randomisiert zugeordnet wurden. Zwischen den Gruppen konnten keine signifikanten Unterschiede hinsichtlich demographischer Merkmale, Ressourcen, Erfahrungen, allgemeinem Fachwissen sowie fachdidaktischem Wissen festgestellt werden. Daher wird von einer homogenen Verteilung der Stichprobenmerkmale zwischen den Gruppen und identischen Gruppenvoraussetzungen ausgegangen. Berechnungen gruppenspezifischer Varianzanalysen, die in den folgenden Kapiteln des Ergebnisteils zu den abhängigen Variablen berichtet werden, sind somit methodisch und statistisch gesehen zulässig.

5.2 Repräsentationsbezogenes Fachwissen und fachdidaktisches Wissen

Zu den repräsentationsbezogenen Variablen zählen das repräsentationsbezogene Fachwissen und fachdidaktische Wissen sowie die repräsentationsbezogenen Überzeugungen. Die zugehörigen Daten wurden sowohl im Prä- als auch Posttest als unabhängige Variablen erfasst und dienen als Grundlage zur Beantwortung der Forschungsfragen 1A und 2. Die Testkennwerte sind in Tabelle 28 dargestellt.

Tabelle 28: Testkennwerte zu den repräsentationsbezogenen Variablen

Variable	Shapiro-Wilk- Test [p]	Schiefte (z -stand.)	Kurtosis (z -stand.)	Reliabilität [α]
Rep. Fachwissen (prä)	< .050	-.43	-.60	.70
Rep. Fachwissen (post)	> .050	.15	-.20	.65
Rep. fachdid. Wissen (prä)	> .050	-.02	-.19	.59
Rep. fachdid. Wissen (post)	> .050	.39	.11	.63
Rep. Überzeugungen (prä)	> .050	-.30	.88	.70
Rep. Überzeugungen (post)	< .050	-1.06	2.60	.81

Die Ergebnisse zum repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen zeigen unterschiedliche Ausprägungen: Während die Ergebnisse zum repräsentationsbezogenen Fachwissen auf dem gleichen Leistungsniveau zwischen 55 und 60 % liegen, fällt das repräsentationsbezogene fachdidaktische Wissen mit Werten von 35 bis 45 % insgesamt deutlich niedriger aus (Tabelle 29).

Die ermittelten Varianzen zwischen den Untersuchungsgruppen sowie den Messzeitpunkten werden im Folgenden dargestellt (Tabelle 29, Abbildung 12 bis Abbildung 15). Das repräsentationsbezogene Fachwissen (Abbildung 12) zeigt im Prätest keine Normalverteilung und wurde daher mittels Kruskal-Wallis-Test auf Varianzen zwischen den Untersuchungsgruppen für jeweils den Prä-

und Posttest untersucht. Im Prätest unterscheidet sich das mittlere repräsentationsbezogene Fachwissen zwischen den drei Untersuchungsgruppen ($\chi^2(2) = 6.92, p = .031$). Anschließend durchgeführte post-hoc-Tests ergeben, dass die Interventionsgruppe 2 im Vergleich zur Kontrollgruppe ein höheres repräsentationsbezogenes Fachwissen aufweist ($z = 2.61, p = .027, \eta^2 = .12$). Im Posttest zeigen sich jedoch keine signifikanten Unterschiede zwischen den Gruppen ($\chi^2(2) = 1.50, p = .471$). Für den Prä-Post-Vergleich wurde ein asymptotischer Wilcoxon-Test durchgeführt. Die Interventionsgruppe 1 ($z = 2.54, p = .011, \eta^2 = .22$) sowie die Kontrollgruppe ($z = 2.99, p = .003, \eta^2 = .29$) zeigen signifikante Verbesserungen, während sich das Fachwissen in der Interventionsgruppe 2 nicht verändert ($z = .04, p = .970$).

Tabelle 29: Vergleich des mittleren relativen repräsentationsbezogenen Fachwissens und fachdidaktischen Wissens

Variable	Untersuchungsgruppe [<i>M, SD</i>]			
	IG1	IG2	KG	Insgesamt
Rep. Fachwissen (prä)	56.59, 17.47	61.26, 14.79	50.37, 17.55	55.88, 17.10
Rep. Fachwissen (post)	63.49, 14.53	61.05, 15.99	59.01, 15.20	61.11, 15.19
Rep. fachdid. Wissen (prä)	41.52, 13.97	38.22, 15.23	36.33, 13.44	38.61, 14.21
Rep. fachdid. Wissen (post)	39.37, 13.34	45.55, 17.43	35.94, 12.60	40.14, 14.93

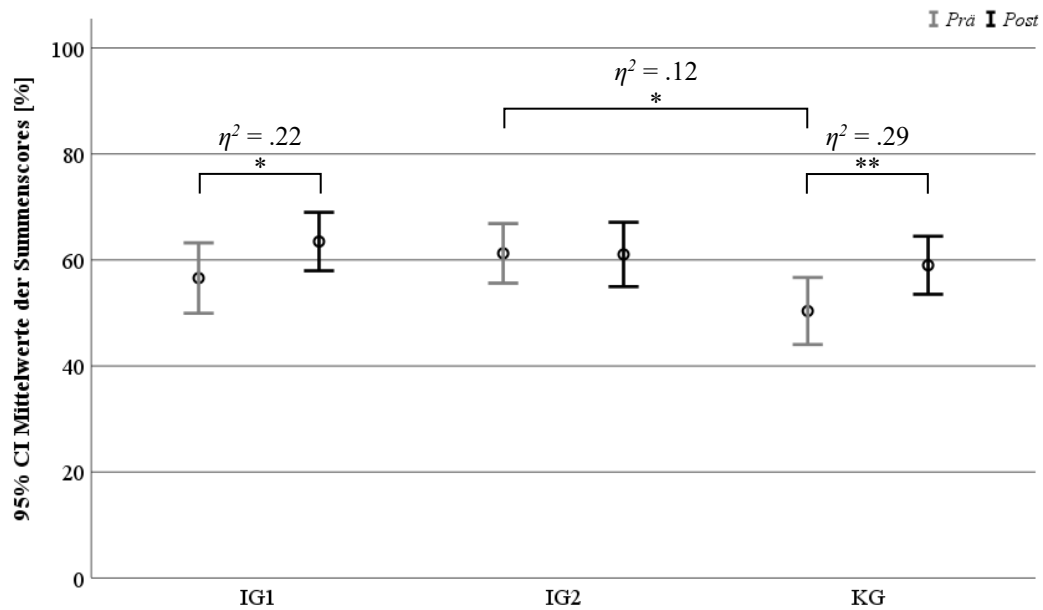


Abbildung 12: Vergleich des mittleren relativen repräsentationsbezogenen Fachwissens

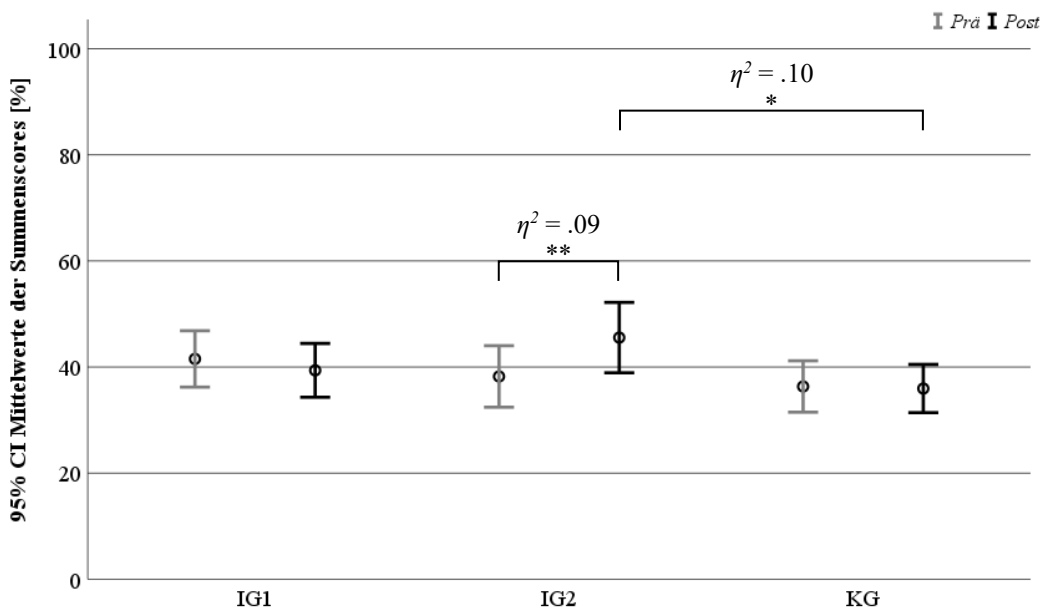


Abbildung 13: Vergleich des mittleren relativen repräsentationsbezogenen fachdidaktischen Wissens

Das repräsentationsbezogene fachdidaktische Wissen zeigt einen anderen Verlauf (Abbildung 13). Die Daten liegen normalverteilt vor, sodass mehrere ANOVA berechnet wurden. Die Untersuchungsgruppen verhalten sich im Prätest gleich ($F(2, 87) = 1.03$, $p = .360$), während im Posttest ein Gruppenunterschied

festgestellt wird ($F(2, 87) = 3.38, p = .039, \eta^2 = .07$). Anschließend durchgeführte Bonferroni-korrigierte post-hoc Tests indizieren einen signifikanten Unterschied im Posttest zwischen der Interventionsgruppe 2 und der Kontrollgruppe ($M_{\text{Diff}} = 9.61, 95\text{-CI} [.50, 18.71], p = .035; t(59) = 2.48, p = .035, \eta^2 = .10$). Eine ANOVA mit Messwiederholung zeigt zudem einen Unterschied in der Performanz der drei Untersuchungsgruppen ($F(2, 87) = 4.84, p = .010, \eta^2 = .10$). Die Interventionsgruppe 2 erwirbt repräsentationsbezogenes Fachwissen signifikant ($t(28) = -3.35, p = .002, \eta^2 = .09$), während bei der Interventionsgruppe 1 ($t(28) = .84, p = .411$) und Kontrollgruppe ($t(31) = .19, p = .848$) kein Lernzuwachs messbar ist.

Bei einer zusammenfassenden Betrachtung der Ergebnisse der Varianzanalysen zum repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen zeigt sich ein Vergleichbarkeitsproblem: Bereits im Prätest gibt es eine Varianz im Fachwissen zwischen der Interventionsgruppe 1 und der Kontrollgruppe, sodass die Lernentwicklungen und Ergebnisse in den Posttests mit den bisherigen absoluten Werten, wie den mittleren relativen Summenscores, statistisch gesehen nicht vergleichbar sind. Zur Lösung dieser Unstimmigkeit werden Residuen der Daten im Fachwissen und fachdidaktischen Wissen herangezogen. Diese stellen die Ergebnisse der beiden Lernzuwächse in bereinigter Form dar, wodurch die unterschiedlichen Voraussetzungen im Prätest ausgeglichen werden.

Die residualen Veränderungen des repräsentationsbezogenen Fachwissens sowie fachdidaktischen Wissens sind normalverteilt (Shapiro-Wilk-Test: $p > .05$) und in der Tabelle 30 dargestellt. Bezogen auf die residuale Veränderung des Fachwissens ($M = .03, SD = .96$) bewegen sich die Residualwerte nahe 0, was darauf hinweist, dass alle drei Untersuchungsgruppen im Vergleich zum erwarteten Lernzuwachs einen ähnlichen Fortschritt erzielt haben (Abbildung 14). Insbesondere zeigt die Interventionsgruppe 1 ($M = .16, SD = .88$) einen moderaten Lernzuwachs. Die Interventionsgruppe 2 ($M = -.16, SD = 1.09$) weist hingegen einen negativen Residualwert auf, was bedeutet, dass in dieser Gruppe der Lernzuwachs unter dem erwarteten Niveau liegt. Die Kontrollgruppe ($M = .07, SD = .91$) zeigt ebenfalls einen geringen, aber positiven Zuwachs. Jedoch ist dieser im Vergleich zu den anderen Gruppen nur in sehr begrenztem Umfang. Insgesamt kann festgestellt werden, dass alle Gruppen einen Zuwachs im reprä-

sentationsbezogenen Fachwissen verzeichnen, auch wenn dieser in der Interventionsgruppe 2 nicht signifikant ausfällt ($F(2, 87) = .85, p = .432$). Die Gruppen sind somit varianzhomogen, sodass alle drei Gruppen gleichermaßen dazulernen.

Tabelle 30: Vergleich der mittleren residualen Veränderung des repräsentationsbezogenen Fachwissens und fachdidaktischen Wissens

Variable	Untersuchungsgruppe [M, SD]			
	IG1	IG2	KG	Insgesamt
Residualer Lernzuwachs im rep. Fachwissen	.16, .88	-.16, 1.09	.07, .91	.03, .96
Residualer Lernzuwachs im rep. fachdid. Wissen	-.22, 1.00	.50, 1.01	-.23, .84	.00, .99

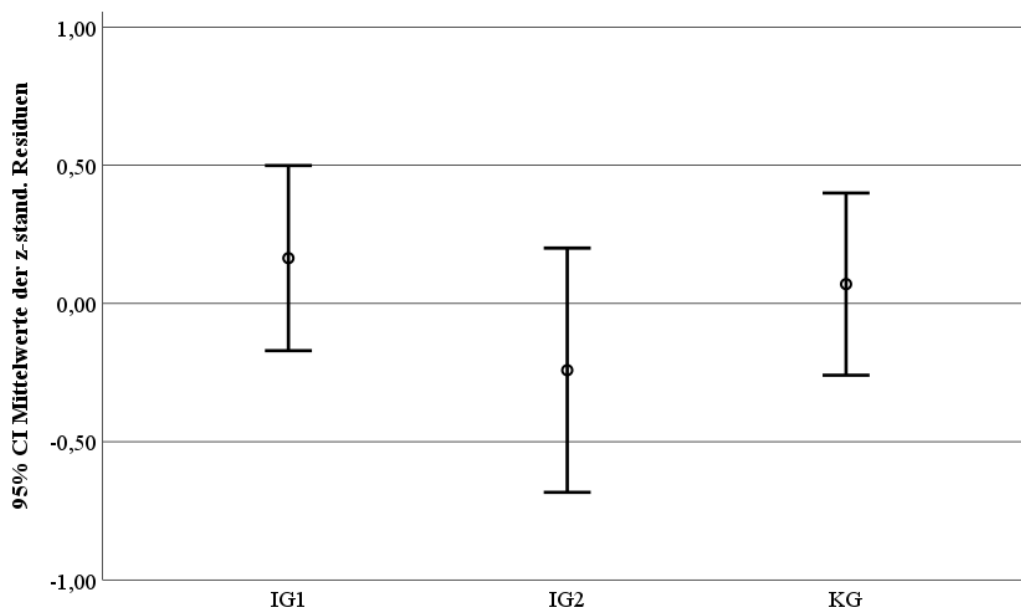


Abbildung 14: Vergleich der mittleren residualen Veränderung des repräsentationsbezogenen Fachwissens

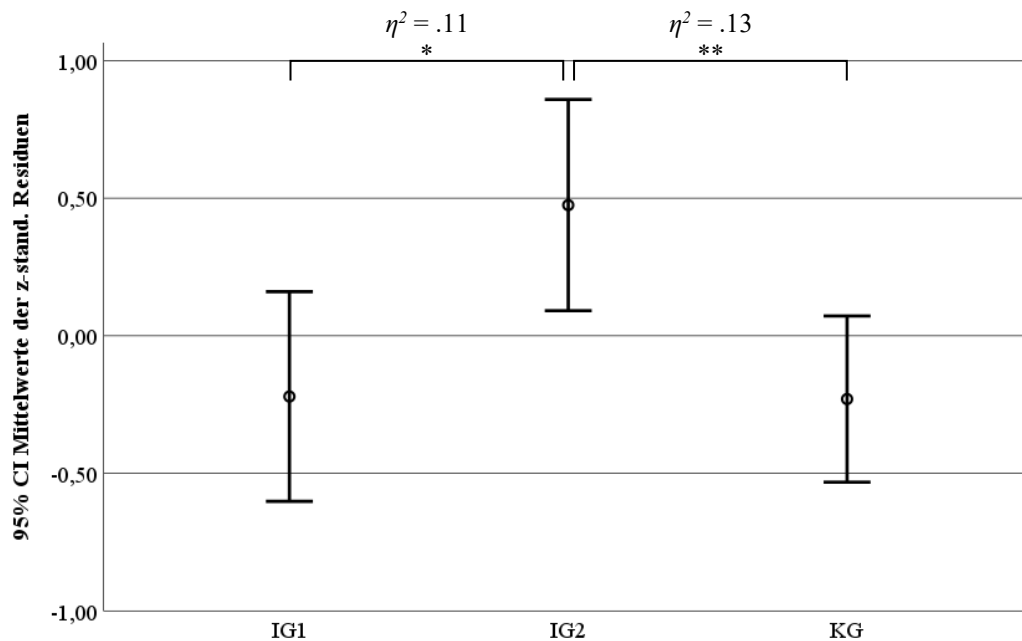


Abbildung 15: Vergleich der mittleren residualen Veränderung des repräsentationsbezogenen fachdidaktischen Wissens

Im Hinblick auf den residualen Lernzuwachs des repräsentationsbezogenen fachdidaktischen Wissens (Abbildung 15) zeigt sich ein differenziertes Bild: Insgesamt ist der Wissenszuwachs im repräsentationsbezogenen fachdidaktischen Wissen über alle Gruppen hinweg sehr unterschiedlich ausgeprägt ($F(2, 87) = 5.67, p = .005, \eta^2 = .07$). Die Interventionsgruppe 1 ($M = -.22, SD = 1.00$) hat im Vergleich zum erwarteten Zuwachs weniger an Wissen gewonnen, während die Interventionsgruppe 2 ($M = .50, SD = 1.01$) einen deutlichen Zuwachs verzeichnet. Die Kontrollgruppe ($M = -.23, SD = .84$) zeigt ebenfalls einen kleineren Lernzuwachs an fachdidaktischem Wissen.

Bonferroni-korrigierte post-hoc Tests zeigen, dass die Interventionsgruppe 2 signifikant mehr Wissen erwirbt, sowohl im Vergleich zur Interventionsgruppe 1 ($M_{\text{Diff}} = .73, 95\%-\text{CI} [.11, 1.34], p = .015; t(56) = 2.63, p = .011, \eta^2 = .11$) als auch zur Kontrollgruppe ($M_{\text{Diff}} = .25, 95\%-\text{CI} [.13, 1.33], p = .011; t(59) = 2.98, p = .004, \eta^2 = .13$). Zusammenfassend erwirbt die Interventionsgruppe 2 im Vergleich zu den anderen beiden Gruppen signifikant und mit mittleren Effektstärken an Wissen.

5.3 Repräsentationsbezogene Überzeugungen

Die repräsentationsbezogenen Überzeugungen der angehenden Lehrkräfte beziehen sich darauf, wie wichtig sie externe Repräsentationen für das Lernen chemischer Inhalte und Konzepte einschätzen und wurden mit einer fünfstufigen Likert-Skala erhoben. Die zugehörigen erfassten Daten dienen als Grundlage der Beantwortung der Forschungsfrage 1B. Bezogen auf die Mittelwerte stufen die angehenden Lehrkräfte Repräsentationen sowohl im Prä- und als auch Posttest insgesamt als wichtig ein (Tabelle 31).

Tabelle 31: Vergleich der mittleren repräsentationsbezogenen Überzeugungen

Variable	Untersuchungsgruppe [M , SD]			
	IG1	IG2	KG	Insgesamt
Rep. Überzeugungen (prä)	2.92, .37	3.11, .31	3.11, .27	3.05, .33
Rep. Überzeugungen (post)	3.02, .44	3.20, .31	3.11, .31	3.11, .36

Anmerkung. Skala: 0 = „gar keine Wichtigkeit von Repräsentationen“ bis 4 = „sehr hohe Wichtigkeit von Repräsentationen“.

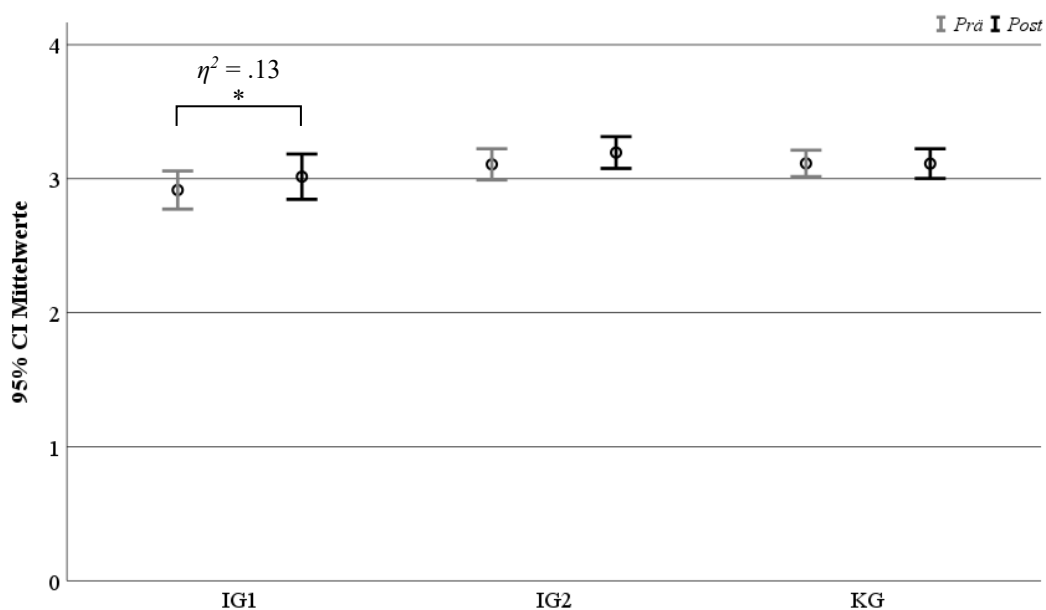


Abbildung 16: Vergleich der mittleren repräsentationsbezogenen Überzeugungen

Die Varianzanalysen wurden mittels Kruskal-Wallis-Test durchgeführt und zeigen sowohl im Prätest ($\chi^2(2) = 4.48, p = .106$) als auch Posttest ($\chi^2(2) = 1.49, p = .475$) keine Unterschiede zwischen den Untersuchungsgruppen. Der Prä-Post-Vergleich mittels Wilcoxon-Tests indiziert eine Verstärkung der Überzeugungen in der Interventionsgruppe 1 ($z = 1.91, p = .046, \eta^2 = .13$). Es sollte jedoch beachtet werden, dass die Überzeugungen der Interventionsgruppe 1 im Prätest deskriptiv niedriger liegen als die der Interventionsgruppe 2 und Kontrollgruppe. Bei der Interventionsgruppe 2 ($z = 1.39, p = .165$) und Kontrollgruppe ($z = -.07, p = .942$) sind keine signifikanten Veränderungen festzustellen. Um die Veränderungen zwischen dem Prä- und Posttest detaillierter zu untersuchen, erfolgt eine Analyse der Residuen (Tabelle 32).

Tabelle 32: Vergleich der residualen Veränderung der repräsentationsbezogenen Überzeugungen

Variable	Untersuchungsgruppe [<i>M, SD</i>]			
	IG1	IG2	KG	Insgesamt
Residuale Veränderung der rep. Überzeugungen	.001, 1.18	.16, .88	-.15, .91	.00, .99

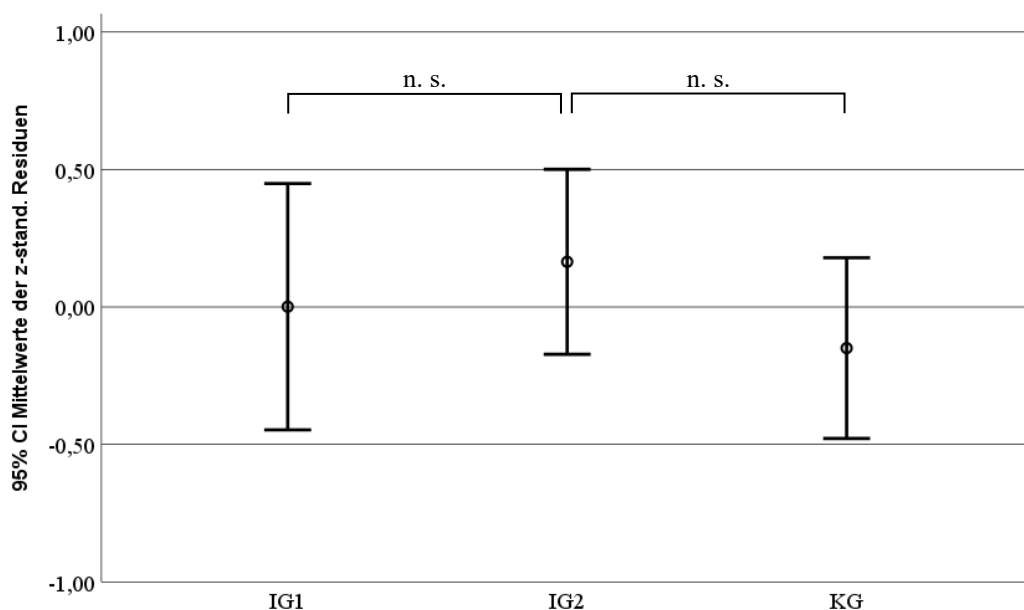


Abbildung 17: Residuale Veränderungen bei den repräsentationsbezogenen Überzeugungen

Die residualen Überzeugungen zu externen Repräsentationen sind nicht normalverteilt (Shapiro-Wilk-Test: $p < .001$), weshalb für die Varianzanalysen ein Kruskal-Wallis-Test durchgeführt wurde. Die mittlere Veränderung der Überzeugungen in den Gruppen liegt nahe Null ($M = .00$, $SD = .99$), was darauf hinweist, dass keine signifikanten Varianzen zwischen den Gruppen bestehen ($\chi^2(2) = 1.75$, $p = .416$). Die Überzeugungen ändern sich in allen drei Gruppen gleichermaßen, wobei die Interventionsgruppe 2 eine geringe positive Veränderung ($M = .16$, $SD = .88$) und die Kontrollgruppe eine minimale negative Veränderung ($M = -0.15$, $SD = 0.91$) aufweist. Es sind somit keine Varianzen festzustellen.

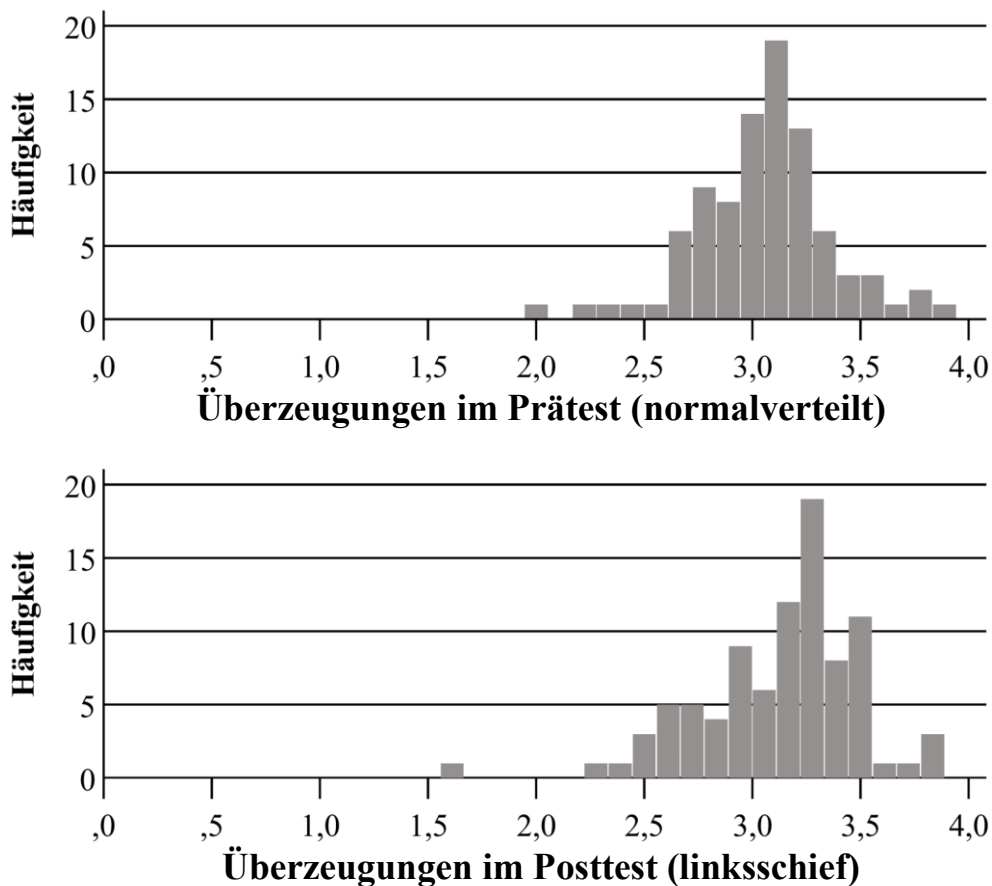
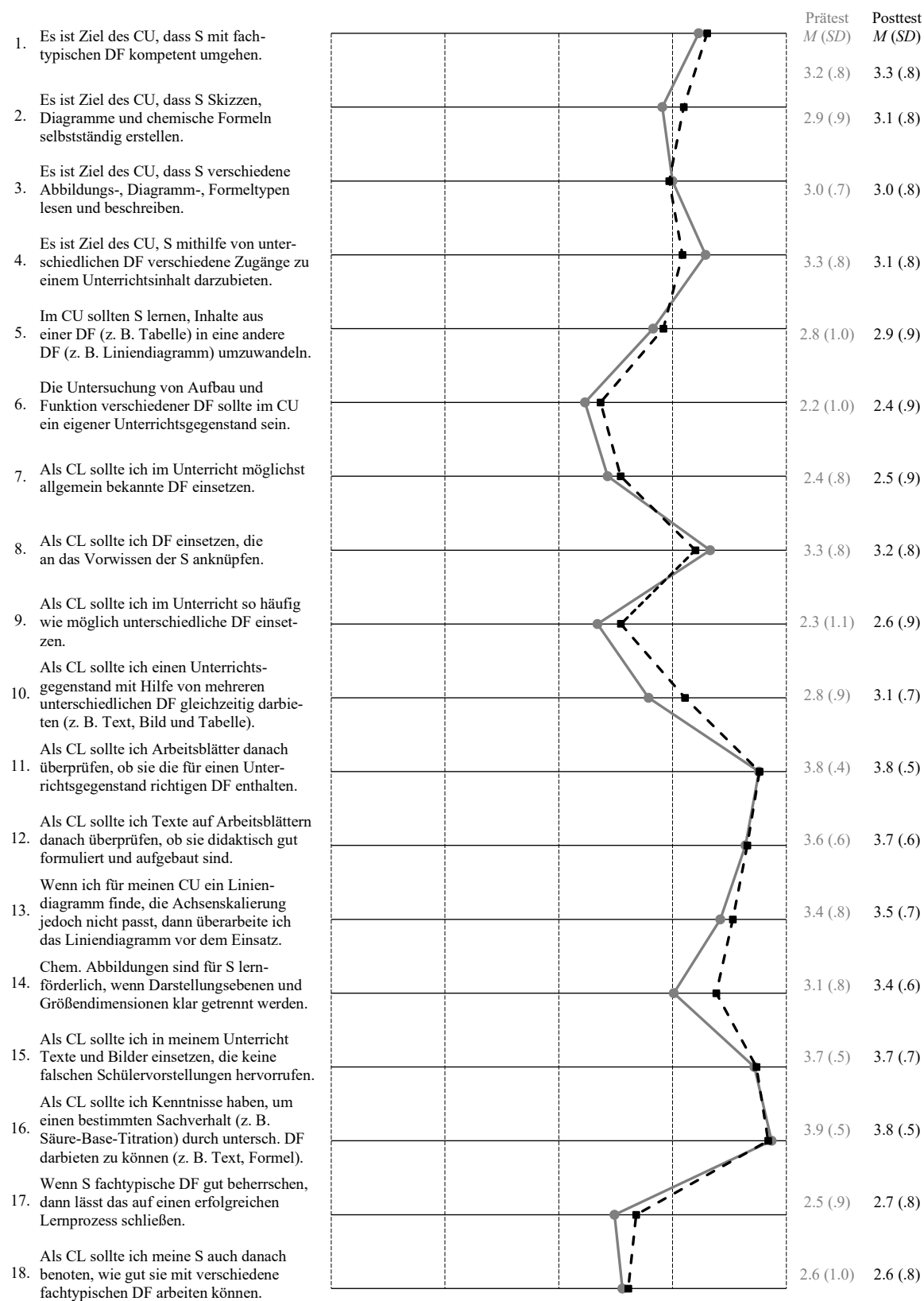


Abbildung 18: Histogramm der mittleren repräsentationsbezogenen Überzeugungen der Gesamtgruppe

Zur vertieften Analyse des Datensatzes wurden zum einen das Histogramm und zum anderen die Profillinien der mittleren Überzeugungen herangezogen. Grund hierfür ist die Normalverteilung, die im Prätest vorzufinden ist und sich im Posttest aufhebt (Tabelle 28). Hierzu wird zunächst in einem Histogramm untersucht, wie sich die mittleren Überzeugungen der einzelnen Individuen aus der Gesamtgruppe verteilen (Abbildung 18). Bei der visuellen Analyse der Histogramme wird auffällig, dass sich die Mittelwerte zu beiden Testzeitpunkten und in allen Untersuchungsgruppen um den Wertebereich 3 (= hohe Wichtigkeit) gruppieren, sich die Häufigkeiten jedoch auf der x-Achse unterschiedlich verteilen. In dem Histogramm des Posttests zeigt sich ein Deckeneffekt, welcher auch durch die Parameter einer Verteilungskurve unterstützt wird. Die Schiefe des Histogramms verändert sich von einer nahezu symmetrischen Verteilung im Prätest (Schiefe = -0.30) zu einer linksschiefen Verteilung im Posttest (Schiefe = -1.06). Darüber hinaus weist die Wölbung bzw. Kurtosis im Prätest einen Wert von $.88$ auf, was auf eine normalgipflige Verteilung hinweist. Im Posttest hingegen ist die Wölbung steilgipfliger (Kurtosis = 2.60), was auf stärkere Peaks und eine konzentrierte Gruppierung der Daten hinweist. Dieser Befund wird durch die Verteilungsfreiheit der residualen Daten weiter gestützt, die ebenfalls eine linksschiefe Verteilung aufweisen (Schiefe = -1.01 , Kurtosis = 1.72).

Die Profillinie in Tabelle 33 bildet die Überzeugungen der angehenden Lehrkräfte in der Gesamtgruppe ab. Betrachtet wird die gesamte Stichprobe, da keine gruppenspezifischen Varianzen in den Überzeugungen der angehenden Lehrkräfte festgestellt wurden. Im Prä-Post-Vergleich zeigt sich, dass sich die Überzeugungen nach der Intervention nicht verändern. Auch lassen sich keine klaren Muster oder Tendenzen erkennen. Es fällt jedoch auf, dass die angehenden Lehrkräfte einige Items als besonders wichtig einstufen, insbesondere in Bezug auf die Aufgaben einer Lehrkraft im Umgang mit Repräsentationen und das Gestalten von Lehr-Lern-Material (Items 16, 11, 15, 12). Diese Items erreichen Mittelwerte von etwa 4 („sehr wichtig“). Andererseits gibt es Items, die als weniger wichtig empfunden werden und Werte von etwa 2.5 („teilweise wichtig“ bis „wichtig“) erreichen, wobei diese Werte dennoch als hoch einzustufen sind. Die angehenden Lehrkräfte finden es am wenigsten wichtig, fachspezifische Repräsentationen im Unterricht nur zu behandeln, weil es sie gibt und sie Teil der Chemie sind (Items 6, 7, 9).

Die Items sind im Graphen gemäß ihrer thematischen Zuordnung zu den Subskalen in Kapitel 4.2.4 geordnet. Eine explorative Faktorenanalyse hat jedoch gezeigt, dass die 18 Items zu den Überzeugungen keine klaren Subskalen abbilden. Das adaptierte Testinstrument misst zwar mit einer hohen Reliabilität insgesamt sehr gut, müsste aber nochmals erprobt und untersucht werden, um mögliche Subskalen eindeutig identifizieren zu können. Eine weiterführende Analyse der Subskalen wird hier jedoch nicht verfolgt, um die benannten Forschungsfragen weiterhin zu fokussieren. Da die Subskalen statistisch nicht bestätigt werden konnten, erscheint die Reihenfolge der Items im Diagramm zufällig. Der genaue Verlauf der Profillinie über die Items wird daher nicht weiter berücksichtigt.

Tabelle 33: Profillinien zu den Überzeugungen der Gesamtstichprobe im Prä-Post-Vergleich


Anmerkung. CL = Chemielehrkraft, CU = Chemieunterricht, DF = Darstellungsformen, S = Schüler:innen.

5.4 Qualität der designten Lehr-Lern-Materialien

Bei den bisher dargestellten Untersuchungsvariablen wurden Ergebnisse berichtet, die zum Prä- und Postmesszeitpunkt bzw. vor und/oder nach der Intervention erfasst wurden. Um die Befunde genauer erklären und ebenso die Forschungsfrage 1C beantworten zu können, werden im Folgenden Prozessdaten dargestellt, d. h. Daten, die während der Intervention bzw. zum Abschluss jeder der drei Sitzungen erhoben wurden. Analysiert wird zunächst die repräsentationsbezogene Gesamtqualität der von den angehenden Lehrkräften designten Lehr-Lern-Materialien. Hierzu werden Gesamtmittelwerte über den gesamten Bewertungsbogen bzw. die acht Qualitätsmerkmale gebildet und für jede Untersuchungsgruppe analysiert. Erst im zweiten Schritt erfolgt die Untersuchung der einzelnen Qualitätsmerkmale sowie die Entwicklung der Untersuchungsgruppen im Interventionsverlauf.

Die Gesamtqualitäten (Abbildung 19) werden jeweils mit der Qualität des Rohmaterials (Kapitel 4.1.4) verglichen, um die Verbesserung des Materials zu verdeutlichen. Allen drei Untersuchungsgruppen gelingt es, aus dem niedrig bewerteten Rohmaterial qualitativ hochwertige Lehr-Lern-Materialien zu gestalten. Dies gelingt auch der Kontrollgruppe, die ohne den Bewertungsbogen arbeitete. Durchgeführte Varianzanalysen mittels Kruskal-Wallis- und Wilcoxon-Tests zeigen, dass es zu allen drei Messzeitpunkten während der Intervention keine Unterschiede zwischen den Untersuchungsgruppen gibt. Die Qualitäten bleiben auch im Verlauf der Intervention gleich und verändern sich nicht. Die Scores liegen aber alle insgesamt bereits auf einem guten Qualitätsniveau.

Im Folgenden werden die Qualitäten der Lehr-Lern-Materialien nach den Qualitätsmerkmalen des Bewertungsbogens unterteilt. Diese umfassen acht Themenblöcke: 1. Passung des Lehr-Lern-Materials zum **Lernziel**, 2. Passung des **Inhaltsumfangs** zur Lerngruppe, 3. Qualität des **Bilds**, 4. Qualität des **Texts**, 5. Passung und Vollständigkeit der **Einleitung**, 6. Korrekte Darstellung der **Repräsentationsebenen**, 7. Korrekte Formulierung und Vollständigkeit der **Aufgabenstellungen** und 8. Positionierung und **Anordnung** der Informationen (Kapitel 4.2.5). Die Merkmale werden zunächst für jede Sitzung einzeln betrachtet und anschließend als Gesamtmittelwert pro Merkmal bezogen auf alle drei Sitzungen.

In Abbildung 20 sind die Mittelwerte für jedes Qualitätsmerkmal in den ersten Sitzungen dargestellt, jeweils für das Rohmaterial sowie für die drei Untersuchungsgruppen. Da die angehenden Lehrkräfte das Lehr-Lern-Material designt und erst anschließend (je nach Untersuchungsgruppe) den Bewertungsbogen gesehen haben, erfolgte bei der Gestaltung diesen Lehr-Lern-Materials noch keine gruppenspezifische Unterweisung. Die Daten können daher als gemeinsame Baseline der Untersuchungsgruppen betrachtet werden. Insgesamt schwanken die Werte im ersten Lehr-Lern-Material zwischen 1.9 und 3.3, was auf eine mittelmäßige bis gute Qualität hindeutet. Die Merkmale Bild und Text wurden in allen Gruppen am besten ausgearbeitet, während alle anderen Werte unter 3 liegen. Auffällig ist, dass die Einleitung in allen Gruppen nicht vollständig auf einem zufriedenstellenden Niveau ausgearbeitet wurde. Gruppenspezifische Varianzen sind in dem ersten Lehr-Lern-Material zu den Merkmalen Lernziel ($\chi^2(2) = 8.2$, $p = .026$; IG1/IG2: $z = 2.87$, $p = .012$, $\eta^2 = .09$) und Text ($\chi^2(2) = 13.00$, $p < .001$; IG1/KG: $z = 3.59$, $p < .001$, $\eta^2 = .21$) zu finden. Auf der Gesamtskala zwischen 0 und 4 sind diese Differenzen allerdings so gering, dass diese nicht weiter beachtet werden.

Beim zweiten Lehr-Lern-Material wurde im Rohmaterial der textliche Teil zur Einleitung komplett ausgelassen. Es wurde bei der Vorbereitung des Rohmaterials davon ausgegangen, dass die Interventionsgruppen die Kriterien zur Einleitung nun kennen und mindestens erkennen, dass ein Kontext im Lehr-Lern-Material vorhanden sein sollte. Auffällig ist hier, dass die Interventionsgruppen den Kontext teilweise hinzufügen, die Kontrollgruppe dies aber nahezu nicht vornimmt ($\chi^2(2) = 8.17$, $p = .02$; IG1/KG: $z = 2.70$, $p = .023$, $\eta^2 = .09$).

Dieser Effekt tritt in der dritten Sitzung nicht mehr auf. Im dritten Rohmaterial wurden zusätzliche Qualitätsmerkmale reduziert, um das neu erlernte Wissen und die neu erlernten Fähigkeiten der angehenden Lehrkräfte besser zu nutzen und das daraus resultierende Verbesserungspotential mehr auszuschöpfen. Alle Qualitätsmerkmale liegen beim Rohmaterial auf einem mangelhaften Niveau vor und werden von den angehenden Lehrkräften aus allen Gruppen auf ein gutes Qualitätsniveau bis auf einen Wert von etwa 3.3 verbessert. Dies gilt für alle Merkmale bis auf die Einleitung. Auch wenn einige Elemente wie Aufgabenstellungen im Rohmaterial vollständig fehlten, wurden diese ergänzt. Auffällig ist, dass dies auch für die Kontrollgruppe zutrifft.

Ein weiterer auffallende Aspekt im dritten Lehr-Lern-Material ist, dass der obere Bereich der Skala von den angehenden Lehrkräften weitestgehend nicht erreicht werden konnte. Gruppenspezifische Varianzen sind messbar bei den Qualitätskriterien Bild ($\chi^2(2) = 13.45$, $p = .00$; IG1/KG $z = 3.09$, $p = .033$, $\eta^2 = .09$; IG2/KG $z = 3.25$, $p = .001$, $\eta^2 = .12$) und Ebenen ($\chi^2(2) = 10.25$, $p = .000$; IG2/KG $z = 3.13$, $p < .001$, $\eta^2 = .11$). Für die Gesamtskala ist der Unterschied sehr klein, sodass nicht verallgemeinert werden kann. Umso deutlicher wird in der dritten Sitzung, dass die Kontrollgruppe im Vergleich zu den beiden Interventionsgruppen eine rückläufige Tendenz aufweist. Sämtliche Mittelwerte der Kontrollgruppe liegen in der dritten Sitzung am niedrigsten.

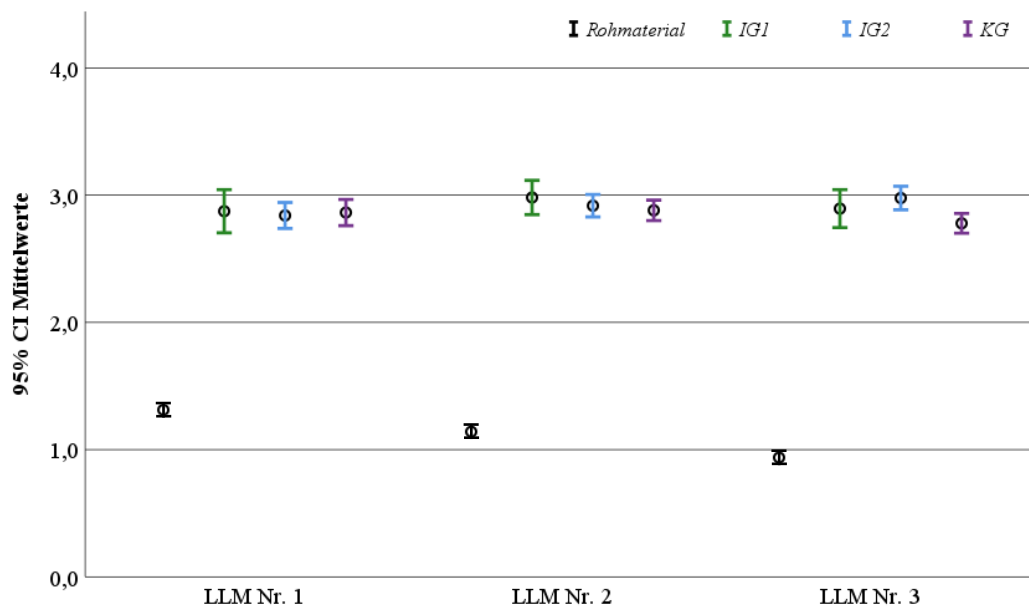


Abbildung 19: Mittlere Gesamtqualität der Lehr-Lern-Materialien der Gesamtgruppe

Anmerkung. LLM = Lehr-Lern-Material. Skala: 0 = „sehr niedrig“ bis 4 = sehr hoch“.

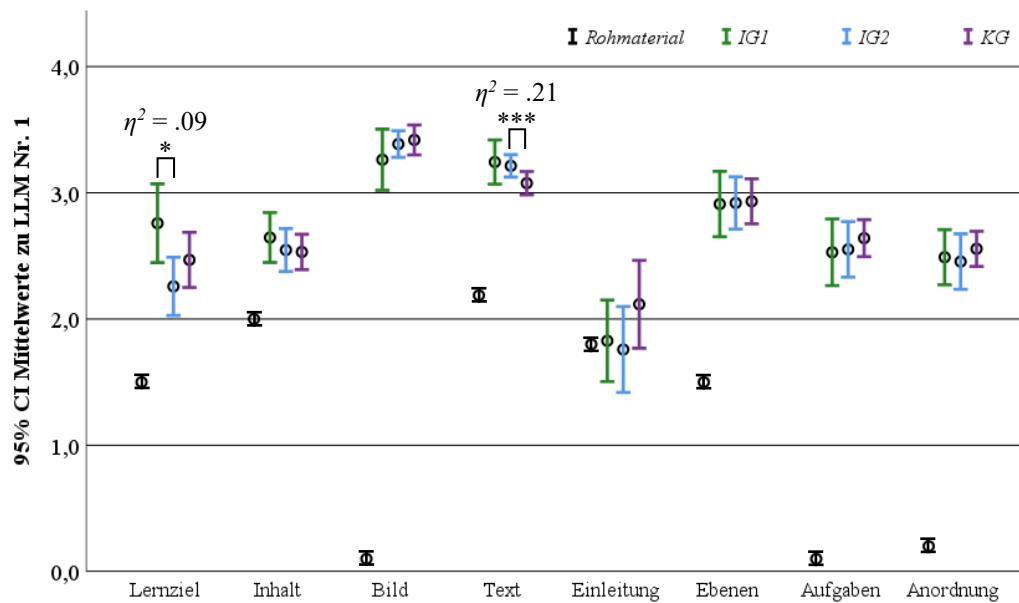


Abbildung 20: Mittlere Qualität des Lehr-Lern-Materialien Nr. 1 der Gesamtgruppe unterteilt nach den Qualitätsmerkmalen des Bewertungsbogens

Anmerkung. LLM = Lehr-Lern-Material. Skala: 0 = „sehr niedrig“ bis 4 = sehr hoch“.

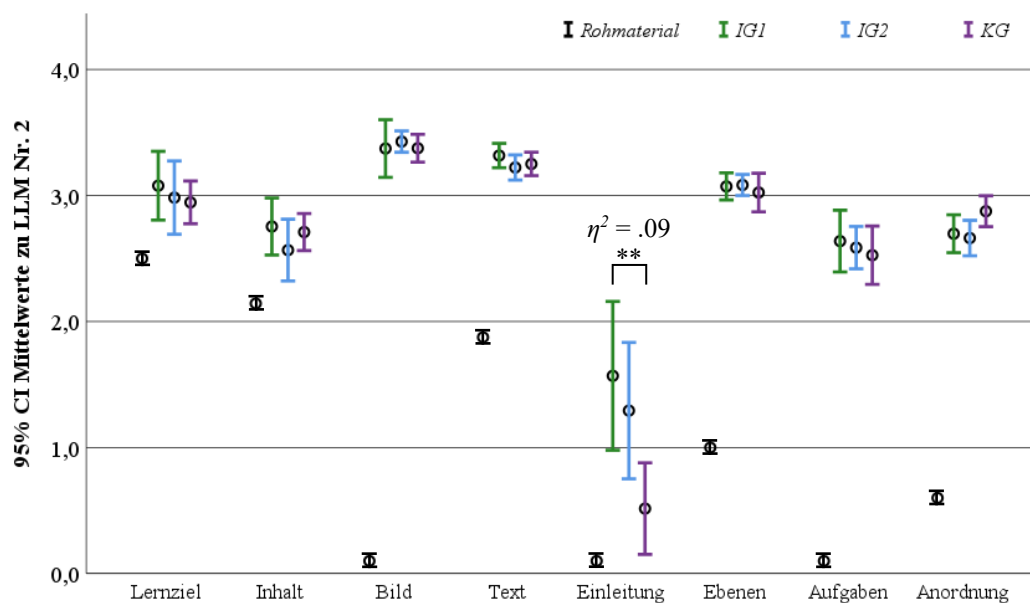


Abbildung 21: Mittlere Qualität der Lehr-Lern-Materialien Nr. 2 der Gesamtgruppe unterteilt nach den Qualitätsmerkmalen des Bewertungsbogens

Anmerkung. LLM = Lehr-Lern-Material. Skala: 0 = „sehr niedrig“ bis 4 = sehr hoch“.

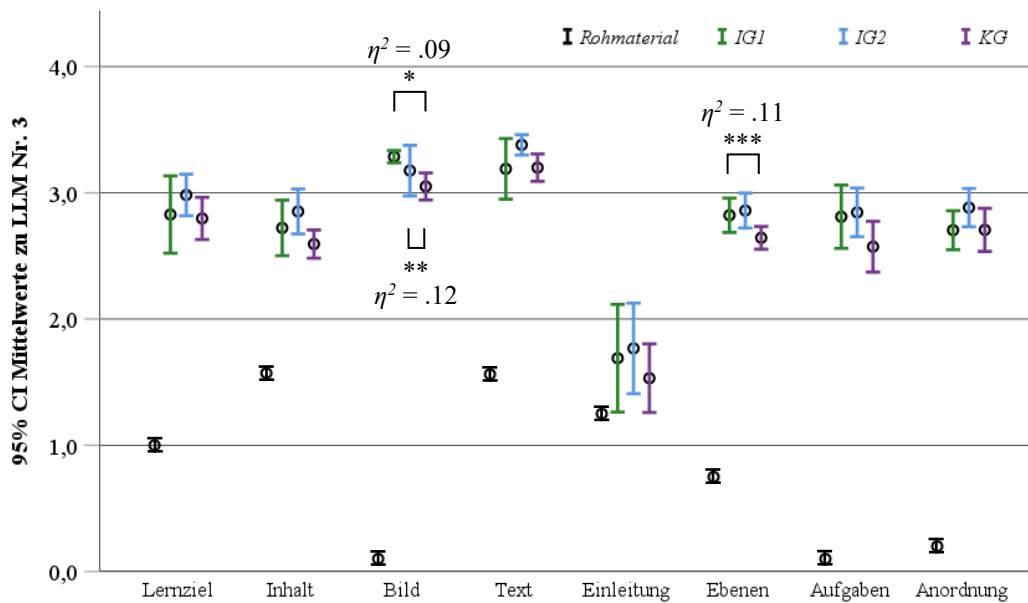


Abbildung 22: Mittlere Qualität des Lehr-Lern-Materials Nr. 3 für die Gesamtgruppe unterteilt nach den Qualitätsmerkmalen des Bewertungsbogens

Anmerkung. LLM = Lehr-Lern-Material. Skala: 0 = „sehr niedrig“ bis 4 = sehr hoch“.

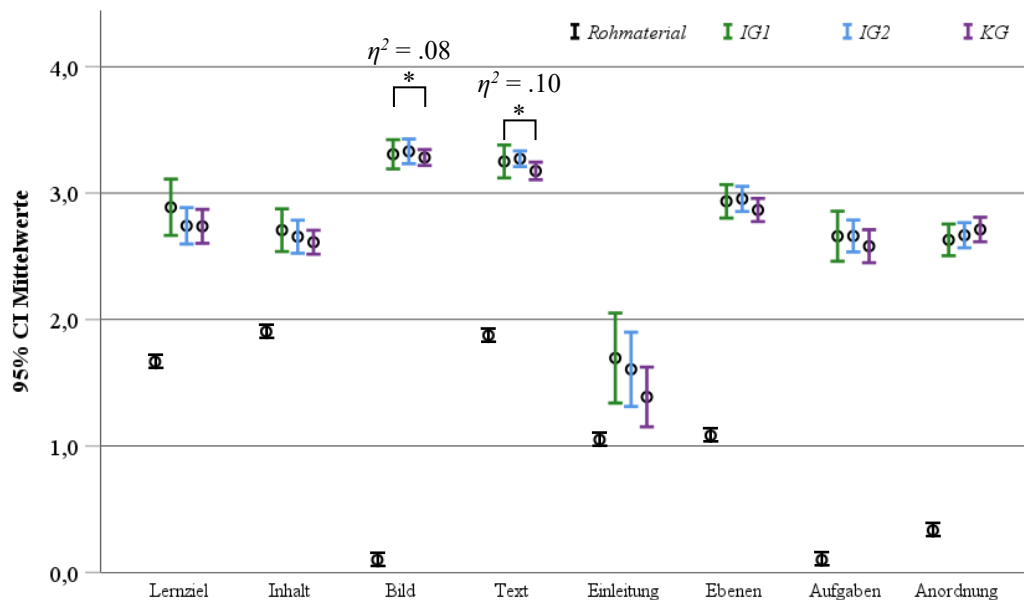


Abbildung 23: Mittlere Qualität der Lehr-Lern-Materialien (Nr. 1, 2 und 3 zusammengefasst) der Gesamtgruppe unterteilt nach den Qualitätsmerkmalen des Bewertungsbogens

Anmerkung. LLM = Lehr-Lern-Material. Skala: 0 = „sehr niedrig“ bis 4 = sehr hoch“.

Zusammenfassend lässt sich unter Betrachtung der Gesamtmittelwerte (Abbildung 23) feststellen, dass alle drei Gruppen bereits vor Beginn der Intervention Vorwissen aufweisen, wie ein Lehr-Lern-Material im Hinblick auf multiple externe Repräsentationen sinnvoll gestaltet sein soll. Ausgehend vom Rohmaterial lässt sich die Qualität in allen Kategorien – mit Ausnahme der Einleitung – im Bereich von etwa 3 einordnen, was einer hohen Qualität entspricht. Dabei scheint es irrelevant zu sein, wie stark oder schwach einzelne Merkmale im Rohmaterial ausgeprägt sind. Die Qualität der Einleitung hingegen bleibt durchgängig auf einem eher niedrigen Niveau. Varianzanalysen zeigen Unterschiede zwischen der Interventionsgruppe 1 und der Kontrollgruppe hinsichtlich der Darstellung von Bild ($\chi^2(2) = 5.94, p = .041; z = 2.18, p = .029, \eta^2 = .08$) und Text ($\chi^2(2) = 6.24, p = .044; z = 2.45, p = .043, \eta^2 = .10$). Diese Unterschiede sind jedoch aufgrund der sehr geringen Effektstärken kaum interpretierbar.

In den bisherigen Untersuchungen zur Qualität der Lehr-Lern-Materialien wurden alle Gruppen gleichzeitig betrachtet, um gruppenspezifische Muster oder Varianzen zu identifizieren. Im Folgenden erfolgt eine getrennte Betrachtung der Untersuchungsgruppen, um deren Entwicklungen innerhalb der einzelnen Sitzungen sowie im Hinblick auf die jeweiligen Qualitätsmerkmale genauer analysieren zu können.

Zunächst wird die Entwicklung der Interventionsgruppe 1 betrachtet, welche sowohl mit internem als auch externem Feedback gearbeitet hat (Abbildung 24, Tabelle 58). Hierbei lassen sich jedoch keine klaren Muster oder Tendenzen erkennen. Auch bei der Interventionsgruppe 2, die ausschließlich mit internem Feedback gearbeitet hat (Abbildung 25, Tabelle 59), sowie der Kontrollgruppe (Abbildung 26) zeigen sich keine deutlichen Verläufe. Auffällig ist jedoch die deutlich höhere Varianz der Mittelwerte innerhalb der Kontrollgruppe. Die Qualität der Lehr-Lern-Materialien schwankt auf Ebene der einzelnen Qualitätsmerkmale erheblich: Teilweise verbessern sich die Ausprägungen (z. B. beim Lernziel), nur um sich im weiteren Verlauf wieder zu verschlechtern. In anderen Fällen zeigen sich kontinuierliche Rückgänge (z. B. beim Bild), oder die Merkmale verändern sich unregelmäßig und stark (z. B. bei der Einleitung).

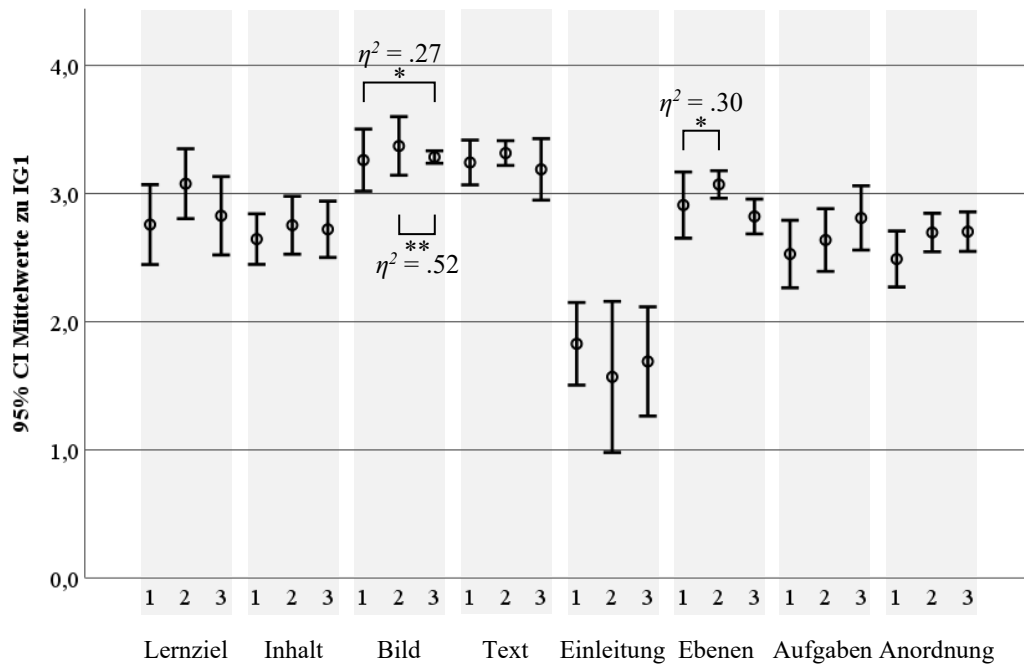


Abbildung 24: Mittlere Qualität der Lehr-Lern-Materialien der Interventionsgruppe 1 unterteilt nach den Sitzungen (Nr. 1, 2 und 3) und Qualitätsmerkmalen

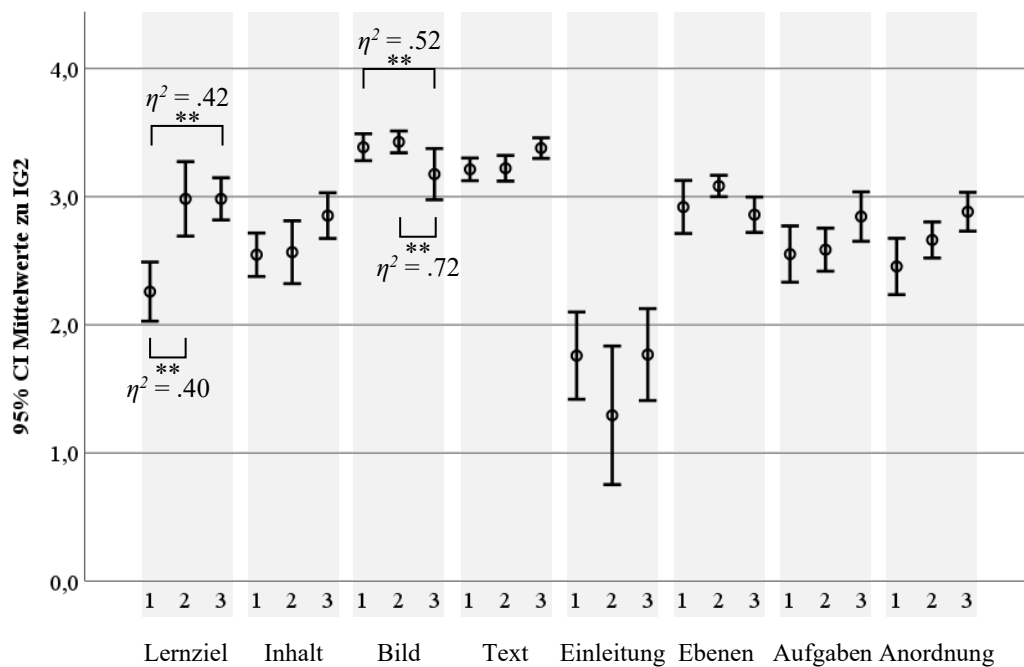


Abbildung 25: Mittlere Qualität der Lehr-Lern-Materialien der Interventionsgruppe 2 unterteilt nach den Sitzungen (Nr. 1, 2 und 3) und Qualitätsmerkmalen

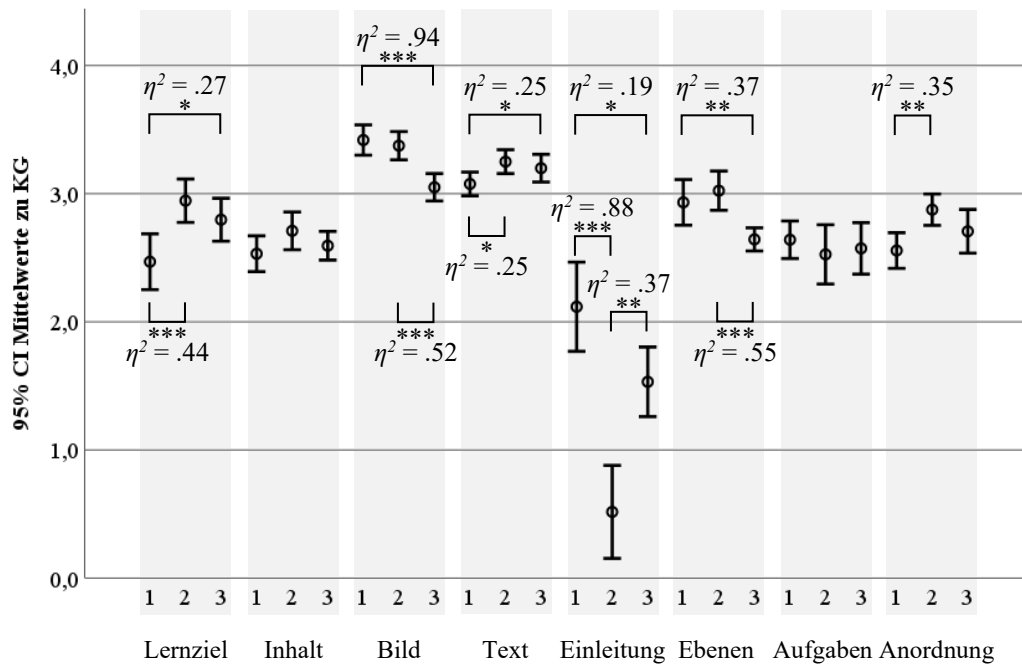


Abbildung 26: Mittlere Qualität der Lehr-Lern-Materialien der Kontrollgruppe unterteilt nach den Sitzungen (Nr. 1, 2 und 3) und Qualitätsmerkmalen

Zusammenfassend lässt sich feststellen, dass die Ausprägungen der Qualitätsmerkmale in den Interventionsgruppen (IG1 und IG2) insgesamt vergleichsweise stabil bleiben, ohne jedoch klare Entwicklungsrichtungen aufzuzeigen. Im Gegensatz dazu zeigt die Kontrollgruppe eine deutlich geringere Stabilität, welche an den erheblichen Schwankungen und unregelmäßigen Verläufen einzelner Qualitätsmerkmale erkennbar ist (Abbildung 26).

5.5 Selbstbewertungsfähigkeit (Qualität des internen Feedbacks)

Die Fähigkeit der angehenden Lehrkräfte, ihr eigenes Lehr-Lern-Material zu bewerten, wurde anhand des Bewertungsbogens erfasst (Kapitel 4.2.6). Da nur die Interventionsgruppen mit dem Bewertungsbogen gearbeitet haben, liegen hier keine Daten aus der Kontrollgruppe vor. Um einerseits die Selbstbewertungsfähigkeit der angehenden Lehrkräfte und andererseits die Eignung der Fünfstufigkeit der Bewertungsskala zu analysieren, werden die Ergebnisse jeweils anhand einer fünfstufigen sowie dreistufigen Skala berichtet.

Die Daten zur Selbstbewertungsfähigkeit (Tabelle 34, Abbildung 27) liegen verteilungsfrei vor (Shapiro-Wilk-Test: $p = .086$, $p = .416$, $p = .444$), sodass für Gruppenvergleiche Mann-Whitney-U-Tests und für Vergleiche zwischen Messzeitpunkten Wilcoxon- oder Friedman-Tests eingesetzt wurden. Anhand eines Gruppenvergleichs über die drei Messzeitpunkte ist zu erkennen, dass sich die Selbstbewertungsfähigkeit beider Gruppen verbessern, es gleichzeitig aber auch einen signifikanten Gruppeneffekt gibt ($F(2, 1.88) = 30.05$, $p < .001$, $\eta^2 = .35$). Anschließend durchgeführte post-hoc Tests zeigen einen auffälligen Befund: In der ersten Sitzung sind die Selbstbewertungsfähigkeiten beider Gruppen nahezu gleich. In der zweiten Sitzung ist ein Gruppeneffekt zunächst nur deskriptiv vorhanden. Die Interventionsgruppe 1 ist minimal besser als die Interventionsgruppe 2. Die Varianz verstärkt sich in der dritten Sitzung und ist mit einer großen Effektstärke signifikant ($t(57) = 3.97$, $p < .001$, $\eta^2 = .22$). Gruppenspezifische post-hoc Analysen zwischen den Messzeitpunkten unterstützen die Ergebnisse: Bei der Interventionsgruppe 1 ist eine signifikante Verbesserung von jeder Sitzung zur darauffolgenden messbar (Nr. 1/2: $t(28) = -3.81$, $p < .001$, $\eta^2 = .11$; Nr. 2/3: $t(28) = -5.12$, $p < .001$, $\eta^2 = .18$; Nr. 1/3: $t(28) = -7.28$, $p < .001$, $\eta^2 = .31$). In der Interventionsgruppe 2 hingegen sind die Verbesserungen nur deskriptiv. Die Signifikanz zeigt sich nur zwischen der ersten und letzten Sitzung ($t(28) = -2.89$, $p = .007$, $\eta^2 = .07$).

Tabelle 34: Vergleich der mittleren Selbstbewertungsfähigkeit

LLM Nr.	Untersuchungsgruppe [<i>M</i> , <i>SD</i>]		
	IG1	IG2	Insgesamt
1	.28, .17	.30, .21	.29, .02
2	.42, .13	.38, .16	.40, .02
3	.62, .13	.45, .19	.54, .02

Anmerkung. Skala: .00 = „niedrige Übereinstimmung/Selbstbewertungsfähigkeit“ bis 1.00 = „hohe Übereinstimmung/Selbstbewertungsfähigkeit“.

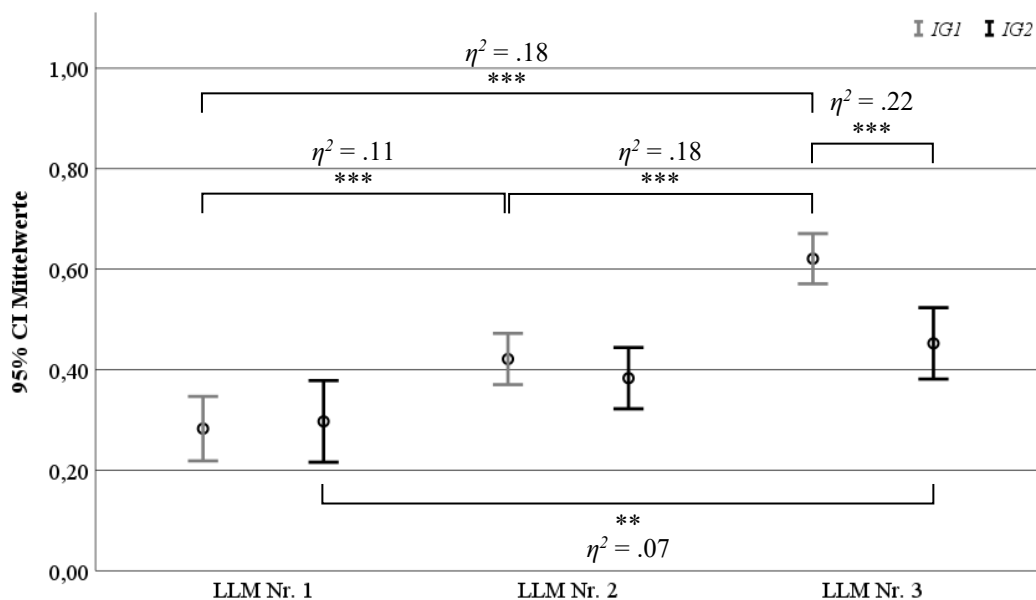


Abbildung 27: Mittlere Selbstbewertungsfähigkeit der Interventionsgruppen 1 und 2 unterteilt nach den Sitzungen (Nr. 1, 2, 3)

Anmerkung. Skala: .00 = „sehr geringe Fähigkeit“ bis 1.00 = „sehr hohe Fähigkeit“.

Im nächsten Schritt wird untersucht, welche Tendenzen die Testpersonen beim Selbstbewerten haben bzw. ob sie sich richtig, zu gut oder zu schlecht einschätzen. Die Skala der in Abbildung 28 dargestellten Ergebnisse ergibt sich aus der absoluten Differenz zwischen dem internen und externen Feedback (4 = zu gute Selbstbewertung, 0 = korrekte/neutrale Selbstbewertung, -4 = zu schlechte Selbstbewertung). Dabei wird der Wert des externen Feedbacks von dem des

internen Feedbacks subtrahiert. Beispiel: Ein gegebenes Item aus dem Bewertungsbogen wird im internen Rating von der angehenden Lehrkraft als 2 eingestuft. Das externe Feedback lautet 3. Als Differenz ergibt sich -1, d. h. die angehende Lehrkraft hat sich leicht schlechter bewertet als sie tatsächlich ist. Aus der Abbildung 28 wird ersichtlich, dass sich die meisten Differenzwerte um die Nulllinie bewegen, was darauf hindeutet, dass die angehenden Lehrkräfte ihre Lehr-Lern-Materialien nur geringfügig besser oder schlechter einschätzen. Insgesamt tendieren die angehenden Lehrkräfte dazu, ihre Abbildungen, Texte und Darstellungen der Repräsentationsebenen positiver zu bewerten, während sie Qualitätsmerkmale, wie den Inhalt, die Einleitung sowie die Aufgabenstellungen negativer bewerten. Signifikant sind diese Ergebnisse für die tendenziell negativere Selbstbewertung der angehenden Lehrkräfte bei der Einleitung ($t(25) = -1.61, p = .027, \eta^2 = .09.$).

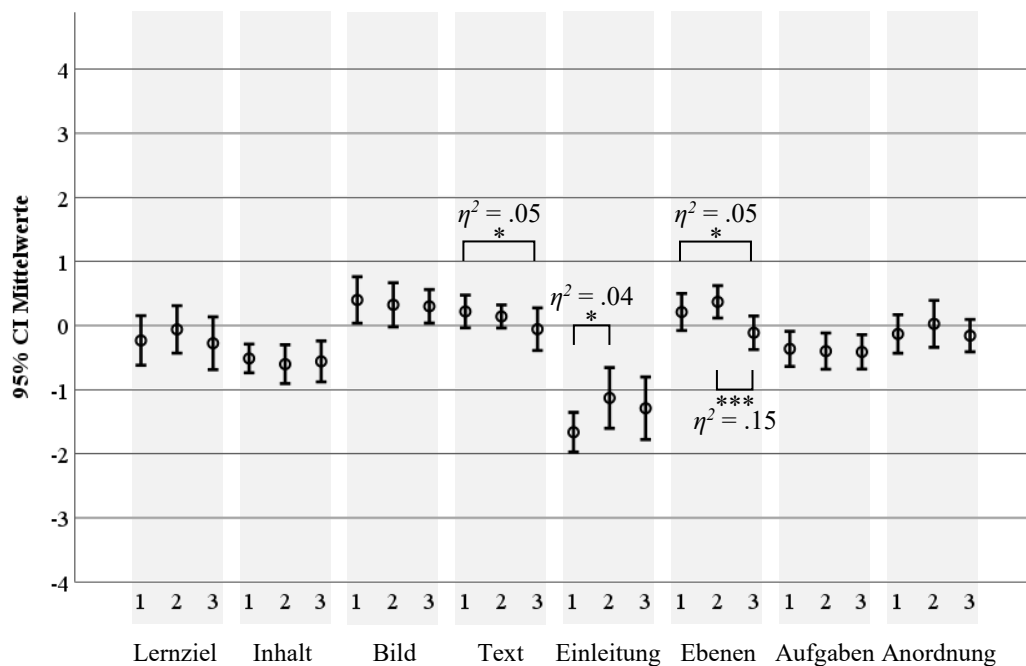


Abbildung 28: Mittlere Bewertungstendenz der Interventionsgruppe 1 bei ihren Lehr-Lern-Materialien unterteilt nach den Sitzungen (Nr. 1, 2 und 3) und Qualitätsmerkmalen

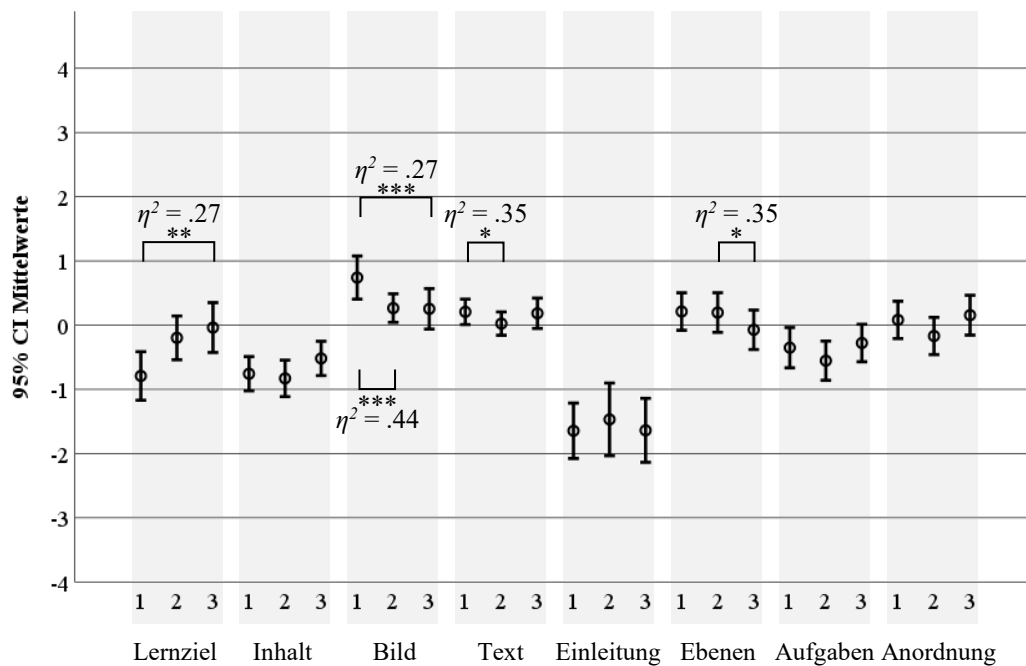


Abbildung 29: Bewertungstendenz der Interventionsgruppe 2 bei ihren Lehr-Lern-Materialien unterteilt nach den Sitzungen (Nr. 1, 2 und 3) und Qualitätsmerkmalen

Anmerkung. Skala: 4.00 = „stark positive Bewertungstendenz“ bis -4.00 = „stark negative Bewertungstendenz“.

Auffällig sind insbesondere die Entwicklungen in den Qualitätsmerkmalen Text, Einleitung und Ebenen. Bei der Entwicklung der Bewertung der Ebenen ist eine ideale und kontinuierliche Verbesserung zu erkennen. Hier zeigt sich eine signifikante Abschwächung der anfänglich zu guten Selbsteinschätzungen im Verlauf der Intervention (Nr. 1/3: $t(28) = 2.55$, $p = .017$, $\eta^2 = .05$; Nr. 2/3: $t(28) = 4.49$, $p < .001$, $\eta^2 = .15$), sodass die Bewertungen am Ende nahezu neutral bzw. korrekt ausfallen und keine Über- oder Unterschätzung mehr vorliegen. Ein vergleichbarer Effekt tritt auch bei der Einleitung auf, wobei hier die signifikante Annäherung an den Nullpunkt bereits zwischen den ersten beiden Sitzungen zu beobachten ist (Nr. 1/2: $t(28) = -2.07$, $p = .011$, $\eta^2 = .04$). Hinsichtlich der Textqualität lässt sich die signifikante Verbesserung der Selbsteinschätzungsfähigkeit zwischen der ersten und letzten Sitzung feststellen (Nr. 1/3: $t(28) = 2.38$, $p = .024$, $\eta^2 = .05$).

Ähnliche Muster zeigen sich bei den Tendenzen der Interventionsgruppe 2 (Abbildung 29). Allerdings schwanken die Differenzwerte hier stärker um die Nulllinie. Generell neigen die angehenden Lehrkräfte der Interventionsgruppe 2 dazu, ihre Bilder, Texte und Darstellungen der Repräsentationsebenen tendenziell besser zu bewerten und Qualitätsmerkmale, wie das Lernziel, den Inhalt, die Einleitung und die Aufgabenstellungen schlechter einzuschätzen. Signifikante Befunde liegen wie in der Interventionsgruppe 1 nur bei der Bewertung der Einleitung vor ($t(28) = -1.61, p = .030, \eta^2 = .09$). Deskriptiv lässt sich eine positive Entwicklung als eine Annäherung an die Nulllinie über alle Qualitätsmerkmale hinweg beobachten. Signifikant sind diese Verbesserungen bei der Selbsteinschätzung in Bezug auf das Lernziel (Nr. 1/3: $t(28) = -3.23, p = .003, \eta^2 = .08$), das Bild (Nr. 1/2: $t(28) = 4.48, p < .001, \eta^2 = .15$; Nr. 1/3: $t(28) = 3.70, p < .001, \eta^2 = .11$), den Text (Nr. 1/2: $t(28) = 2.55, p = .017, \eta^2 = .05$) sowie die Repräsentationsebenen (Nr. 2/3: $t(28) = 2.93, p = .007, \eta^2 = .07$).

5.5.1 Analyse der Eignung der fünfstufigen Skalierung zur Selbsteinschätzung

Im letzten Analyseschritt wird die methodische Eignung der Fünfstufigkeit der Bewertungsskala untersucht und Ergebnisse zu der Selbsteinschätzungsfähigkeit anhand einer dreistufigen Skala dargestellt (Tabelle 35, Abbildung 30). So soll untersucht werden, ob die Selbsteinschätzungsfähigkeit der angehenden Lehrkräfte höher ausfällt, wenn es im zustimmenden bzw. ablehnenden Bereich der Skala jeweils nur eine Ausprägung gibt.

In der Abbildung 30 werden die mittleren Gamma-Werte zur Selbsteinschätzungsfähigkeit jeweils auf einer fünf- und dreistufigen Skala dargestellt. Hierbei zeigt sich ein Anstieg der Mittelwerte über die drei Lehr-Lern-Materialien und Interventionsgruppen hinweg. Allerdings sind die Standardabweichungen relativ hoch. Darüber hinaus steigen die Gamma-Werte der Interventionsgruppe 1 bei einer dreistufigen Skala im Vergleich zur fünfstufigen Skala an. Dies betrifft das zweite ($t(28) = -2.46, p = .020, \eta^2 = .05$) und dritte Lehr-Lern-Material ($t(28) = -2.92, p = .007, \eta^2 = .07$). In der Interventionsgruppe 2 hingegen ergeben sich keine signifikanten Unterschiede.

Tabelle 35: Vergleich der mittleren Selbstbewertungsfähigkeit anhand einer dreistufigen Likert-Skala

LLM Nr.	Untersuchungsgruppe [<i>M</i> , <i>SD</i>]		
	IG1	IG2	Insgesamt
1	.32, .31	.33, .35	.32, .04
2	.52, .22	.41, .46	.46, .05
3	.72, .14	.49, .30	.60, .03

Anmerkung. Skala: .00 = „sehr geringe Fähigkeit und Qualität des internen Feedbacks“ bis 1.00 = „sehr hohe Fähigkeit und Qualität des internen Feedbacks“.

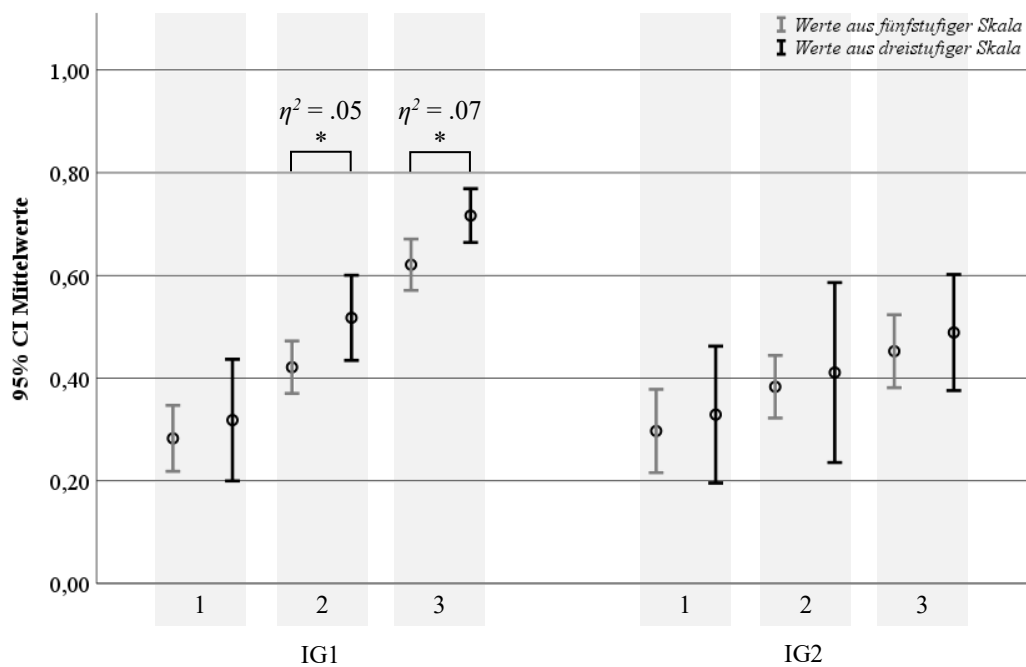


Abbildung 30: Mittlere Selbstbewertungsfähigkeit unterteilt nach den Sitzungen (Nr. 1, 2, 3) und jeweils ausgewertet mit einer fünf- und dreistufigen Likert-Skala

Anmerkung. Skala: .00 = „sehr geringe Fähigkeit und Qualität des internen Feedbacks“ bis 1.00 = „sehr hohe Fähigkeit und Qualität des internen Feedbacks“.

5.5.2 Wahrgenommener Cognitive Load während der Feedbackphasen

Zwei weitere Variablen, die zum Schluss jeder Interventionssitzung gemessen worden sind und deren Ergebnisse zu allen drei Sitzungen vorliegen, sind der wahrgenommene Cognitive Load sowie die Anstrengungsbereitschaft beim Arbeiten mit dem Bewertungsbogen. Auch hier liegen keine Daten zu der Kontrollgruppe vor, da diese während der Intervention ohne Bewertungsbogen gearbeitet hat.

Tabelle 36: Testkennwerte zum Cognitive Load und zur Anstrengungsbereitschaft

Variable	Shapiro-Wilk-Test [p]	Schiefe (z-stand.)	Kurtosis (z-stand.)	Reliabilität [α]
<i>Cognitive Load</i>				
1. Sitzung	< .050	.10	.50	.70
2. Sitzung	< .050	.51	.07	.65
3. Sitzung	< .050	.81	.55	.59
<i>Anstrengungsbereitschaft</i>				
1. Sitzung	< .050	-1.18	1.11	.63
2. Sitzung	< .050	-.38	-1.08	.70
3. Sitzung	< .050	-.51	-.60	.81

Wie in der Tabelle 36 zu erkennen ist, liegen die Daten zu beiden Variablen verteilungsfrei vor. Die Reliabilitäten liegen insgesamt in einem akzeptablen Bereich. Auffällig sind die Werte zur Schiefe und Kurtosis, die auf eine Verschiebung der Verteilung hinweisen. Bei genauerer Betrachtung der Histogramme (Abbildung 31), differenziert nach den einzelnen Variablen und Messzeitpunkten, zeigt sich für den wahrgenommenen Cognitive Load eine Verschiebung des Peaks nach links. Bei einem Wert von 1.00 ist durchweg ein Bodeneffekt zu sehen. Bei der wahrgenommenen Anstrengungsbereitschaft hingegen ist bereits von Beginn an ein Deckeneffekt erkennbar, da sich die Mittelwerte der Testpersonen im oberen bzw. rechten Bereich konzentrieren. In der zweiten und dritten Sitzung vermindert sich dieser Deckeneffekt.

In Abbildung 32 sind die gruppenspezifischen Entwicklungen des Cognitive Loads dargestellt. Zunächst ist ein mittelstarker Cognitive Load zu erkennen,

der während der Intervention abnimmt. Zu allen Messzeitpunkten ist der Cognitive Load der Interventionsgruppe 1 höher als der der Interventionsgruppe 2 (Nr. 1: $U = 294.50$, $Z = -2.02$, $p = .043$, $\eta^2 = .07$; Nr. 2: $U = 302.00$, $Z = -1.89$, $p = .049$, $\eta^2 = .06$; Nr. 3: $U = 268.00$, $Z = -2.41$, $p = .016$, $\eta^2 = .01$). Außerdem nimmt der Cognitive Load in der Interventionsgruppe 2 kontinuierlich signifikant ab ($\chi^2(2) = 25.640$, $p < .001$; Nr. 1/2: $Z = -2.02$, $p = .043$, $\eta^2 = .14$; Nr. 2/3: $Z = -2.90$, $p = .004$, $\eta^2 = .29$; Nr. 1/3: $Z = -3.77$, $p < .001$, $\eta^2 = .49$), während in der Interventionsgruppe 1 eine signifikante Abnahme nur nach der ersten Sitzung zu finden ist ($\chi^2(2) = 14.00$, $p < .001$; Nr. 1/2: $Z = -2.64$, $p = .008$, $\eta^2 = .24$; Nr. 1/3: $Z = -3.04$, $p = .002$, $\eta^2 = .32$).

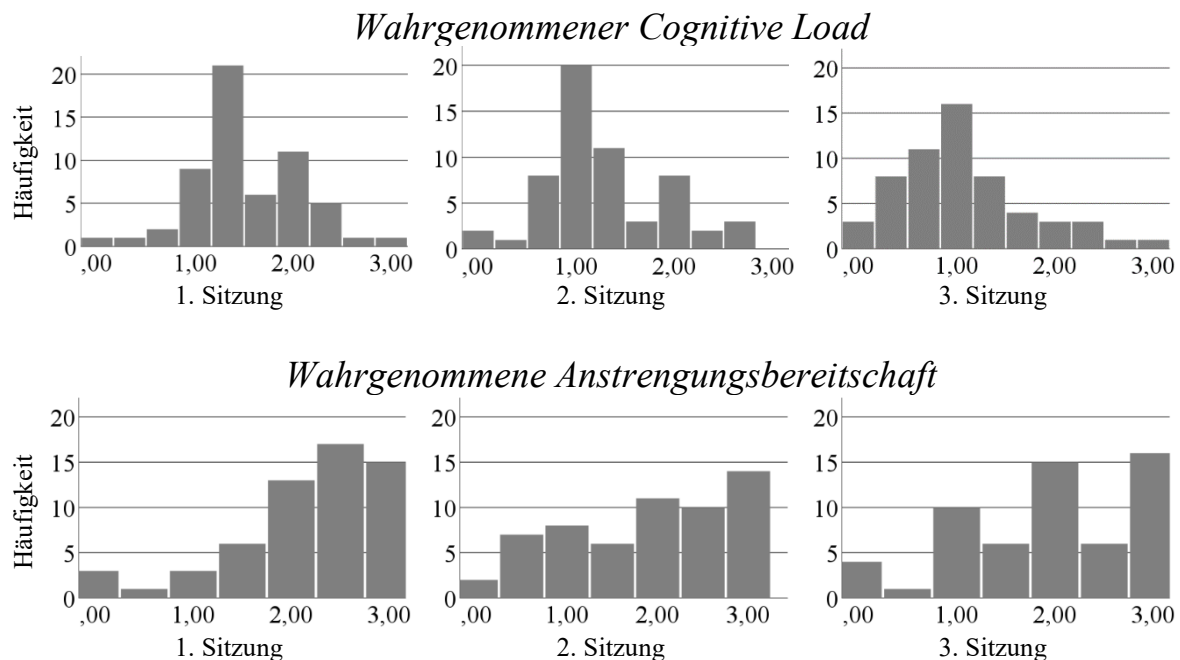


Abbildung 31: Histogramm des Cognitive Loads sowie der Anstrengungsbereitschaft in den Interventionsgruppen 1 und 2, unterteilt nach den Sitzungen

Anmerkung. Skala: .00 = „sehr geringe Ausprägung“ bis 3.00 = „sehr hohe Ausprägung“.

Tabelle 37: Vergleich des mittleren Cognitive Loads

LLM Nr.	Untersuchungsgruppe [<i>M, SD</i>]		
	IG1	IG2	Insgesamt
1	1.64, .58	1.39, .55	1.52, .56
2	1.38, .55	1.17, .66	1.28, .61
3	1.24, .63	.90, .67	1.07, .67

Anmerkung. Skala: .00 = „sehr geringer Cognitive Load“ bis 3.00 = „sehr hoher Cognitive Load“.

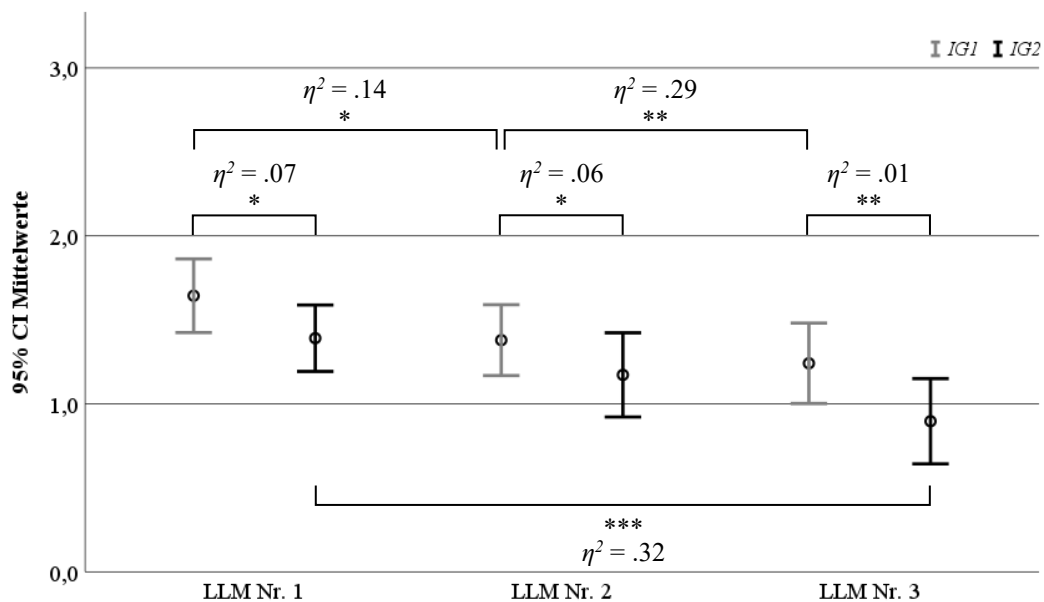


Abbildung 32: Mittlerer wahrgenommener Cognitive Load in den Interventionsgruppen 1 und 2 während der Feedbackphasen

Anmerkung. Skala: .00 = „sehr geringer Cognitive Load“ bis 3.00 = „sehr hoher Cognitive Load“.

5.5.3 Wahrgenommene Anstrengungsbereitschaft während der Feedbackphasen

Die Anstrengungsbereitschaft der angehenden Lehrkräfte gibt an, wie viel Mühe sie sich laut Selbstauskunft beim Ausfüllen des Bewertungsbogens gegeben haben und wie gewissenhaft dies erfolgte (Tabelle 38 und Abbildung 33).

Tabelle 38: Vergleich der mittleren Anstrengungsbereitschaft

LLM Nr.	Untersuchungsgruppe [<i>M</i> , <i>SD</i>]		
	IG1	IG2	Insgesamt
1	2.24, .71	2.10, .88	2.17, .80
2	1.90, .99	1.88, .87	1.89, .93
3	1.76, .95	2.12, .84	1.94, .91

Anmerkung. Skala: .00 = „sehr geringe Anstrengung“ bis 3.00 = „sehr hohe Anstrengung“.

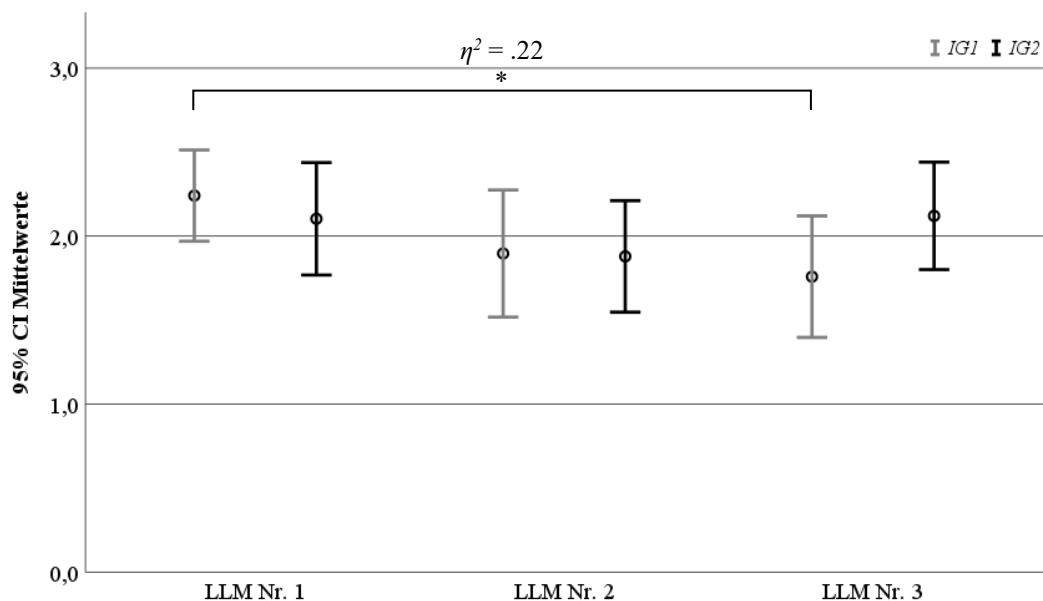


Abbildung 33: Mittlere wahrgenommene Anstrengungsbereitschaft in den Interventionsgruppe 1 und 2 während der Feedbackphasen

Anmerkung. Skala: .00 = „sehr geringe Anstrengung“ bis 3.00 = „sehr hohe Anstrengung“.

Insgesamt liegt die Anstrengungsbereitschaft in einem mittleren bis hohen Bereich. Gruppenspezifische Unterschiede pro Messzeitpunkt sind nicht vorzufinden. Varianzanalysen mit Messwiederholung zeigen allerdings in der Interventionsgruppe 1 eine Abnahme der Anstrengungsbereitschaft ($\chi^2(2) = 7.69, p = .021$; Nr. 1/3: $Z = -2.55, p = .011, \eta^2 = .22$).

5.6 Zusammenhangsanalysen zu den Stichprobenmerkmalen

5.6.1 Varianzen in den demographischen Daten der Gesamtgruppe

Zur Beantwortung von Forschungsfrage 2 wurde explorativ über Varianzanalysen untersucht, ob die Stichprobenmerkmale von angehenden Lehrkräften einen Einfluss auf ihre Ressourcen, Fähigkeiten und ihr Wissen haben. Als ausschlaggebende Stichprobenmerkmale stellen sich hier die Lehramtsform sowie die Art des Hochschulabschlusses heraus (Tabelle 39 bis Tabelle 42).

Bei der Lehramtsform wird zwischen angehenden Lehrkräften unterschieden, die ihren Vorbereitungsdienst für Haupt-, Real-, Sekundar- und Gesamtschulen absolvieren bzw. nur in der Sekundarstufe I unterrichten (HRSGe) und angehenden Lehrkräften, die ihren Vorbereitungsdienst für Gymnasien und Gesamtschulen leisten bzw. somit sowohl in der Sekundarstufe I als auch II eingesetzt werden (GymGe). Bei den angehenden Lehrkräften für Gymnasien und Gesamtschulen zeigen sich nicht nur höhere Werte für die Belastungen und Bewältigungsstrategien, sondern auch ein insgesamt höheres Fachwissen, sowohl im allgemeinen als repräsentationsbezogenen Bereich. Ein umgekehrter Effekt zeigt sich jedoch bei der Selbstwirksamkeitserwartung. Hier erwarten die angehenden Lehrkräfte der Sekundarstufe I eine höhere Wirksamkeit von ihren unterrichtspraktischen Bemühungen.

Tabelle 39: Signifikante Varianzanalysen zwischen angehenden Lehrkräften unterschiedlicher Lehramtsformen

Variable	Lehramtsform [<i>M</i> , <i>SD</i>]		Ergebnis der Varianzanalyse
	HRSGe <i>n</i> = 35	GymGe <i>n</i> = 55	
Zeitdruck	1.50, .52	1.77, .54	$F(1, 89) = 1.54, p < .022, \eta^2 = .06$
Elaborationsmuster	1.51, .47	1.76, .49	$F(1, 89) = 5.84, p = .018, \eta^2 = .06$
Soziale Ressourcen	1.98, .44	2.21, .48	$F(1, 89) = 5.49, p = .021, \eta^2 = .06$
SWE	3.64, .50	3.42, .49	$F(1, 89) = 4.12, p = .045, \eta^2 = .05$
Allg. Fachwissen (prä)	51.24, 18.70	62.06, 17.26	$F(1, 89) = 7.88, p = .006, \eta^2 = .08$
Allg. Fachwissen (post)	56.57, 13.85	64.85, 19.05	$F(1, 89) = 4.94, p = .029, \eta^2 = .05$
Rep. Fachwissen (prä)	51.43, 17.20	58.72, 16.58	$F(1, 89) = 4.01, p = .048, \eta^2 = .04$
Rep. Fachwissen (post)	54.29, 11.77	65.45, 18.61	$F(1, 89) = 13.14, p < .001, \eta^2 = .13$

Anmerkung. SWE = Selbstwirksamkeitserwartung an den eigenen Unterricht.

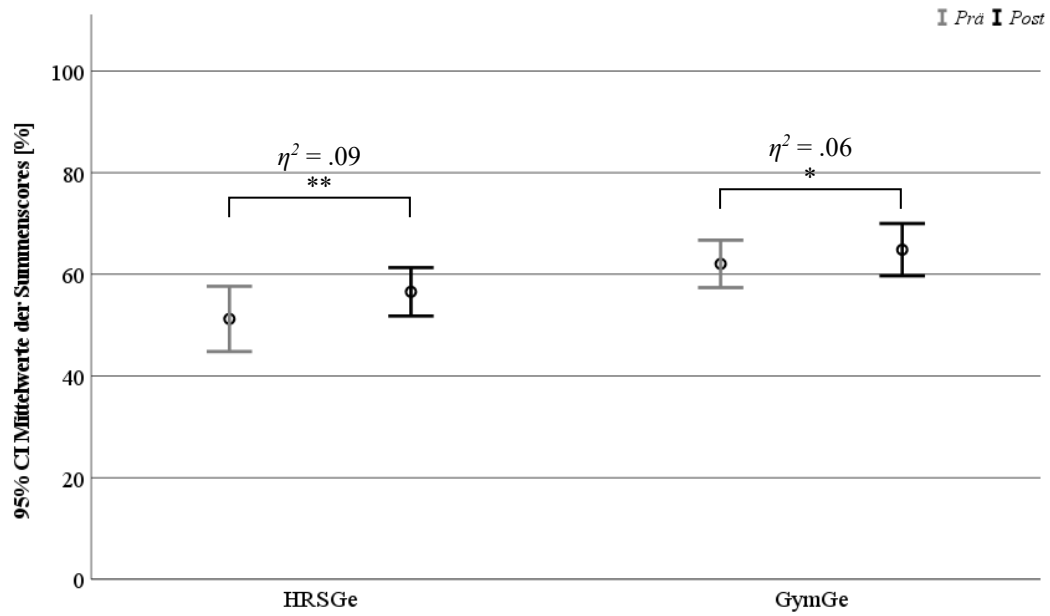


Abbildung 34: Vergleich des mittleren relativen allgemeinen Fachwissens unterteilt nach Lehramtsformen

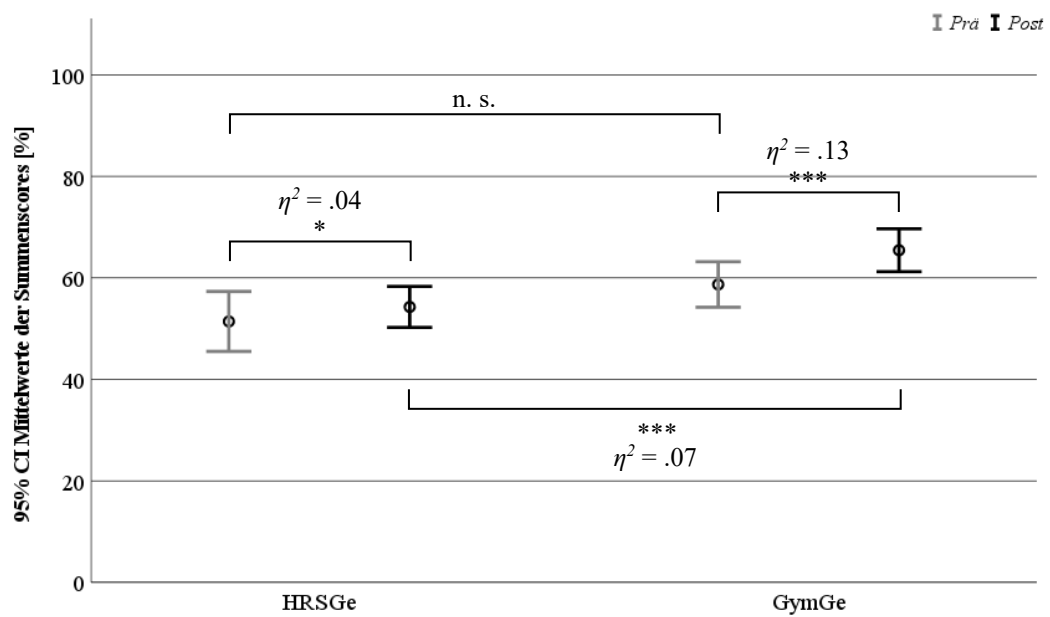


Abbildung 35: Vergleich des mittleren relativen repräsentationsbezogenen Fachwissens unterschiedlicher Lehramtsformen

Im Folgenden werden die lehramtsspezifischen Unterschiede genauer betrachtet, da sie für beide Fachwissensarten und Messzeitpunkte vorliegen und diese auf einen erheblichen Unterschied hindeuten. Die Varianzen im allgemeinen Fachwissen sind in der Abbildung 34 visualisiert. Da es sich beim allgemeinen Fachwissen um eine Kontrollvariable handelt, ist wie erwartet in beiden Untersuchungsgruppen kein Lernzuwachs messbar. Dennoch zeigen sich andere gruppenspezifische Unterschiede ($F(1, 88) = 5.83, p = .018, \eta^2 = .06$). So verfügen die angehenden Lehrkräfte der Sekundarstufe I und II (GymGe) zu beiden Messzeitpunkten über ein höheres allgemeines Fachwissen als die angehenden Lehrkräfte der Sekundarstufe I (Prä: $t(88) = -2.81, p = .006, \eta^2 = .09$; Post $t(88) = -2.22, p = .029, \eta^2 = .06$).

Vergleichbare Ergebnisse zeigen sich auch für das repräsentationsbezogene Fachwissen (Abbildung 35). Auch hier erzielen die angehenden Lehrkräfte der Sekundarstufe I und II insgesamt bessere Leistungen als die angehenden Lehrkräfte der Sekundarstufe I ($F(1, 88) = 9.46, p = .003, \eta^2 = .097$). Der Leistungsunterschied zwischen den Untersuchungsgruppen ist zu beiden Messzeitpunkten signifikant (Prä: $t(88) = -2.00, p = .048, \eta^2 = .04$; Post $t(88) = -3.63, p < .001, \eta^2 = .13$). Auffällig ist zudem, dass die angehenden Lehrkräfte der Sekundarstufe I und II (unabhängig von ihrer Gruppenzugehörigkeit) während der Intervention einen Lernzuwachs aufweisen ($t(54) = -4.09, p < .001, \eta^2 = .07$) und die der anderen Lehramtsform nicht.

Bezogen auf die Art des Hochschulabschlusses, mit dem Vorbereitungsdienst angetreten wurde, zeigen sich Unterschiede zwischen angehenden Lehrkräften, die ein Studium mit Lehramtsoption (Master of Education oder Erstes Staatsexamen) und solchen ohne Lehramtsoption (z. B. Master of Arts, Master of Science, Diplom) erworben haben (Tabelle 40). Die Gruppen unterscheiden sich nicht in ihrem Wissen, sondern in den Belastungen und Bewältigungsstrategien. Außerdem kovariieren die Ergebnisse zum Hochschulabschluss mit dem Alter der Testpersonen. Wird die Gesamtgruppe nach Altersgruppen eingeteilt, ergeben sich ähnliche Ergebnisse wie bei der Differenzierung nach Abschlussart. Dies liegt beispielsweise daran, dass angehende Lehrkräfte eines Studiums ohne Lehramtsoption zu den höheren Altersgruppen gehören. Zudem sind die angehenden Lehrkräfte mit Lehramtsoption signifikant jünger als diejenigen mit einem Studium ohne Lehramtsoption.

Tabelle 40: Signifikante Varianzanalysen zwischen angehenden Lehrkräften unterschiedlicher Hochschulabschlüsse

Variable	Abschluss [<i>M</i> , <i>SD</i>]		Ergebnis der Varianzanalyse
	mit Lehramts- option <i>n</i> = 60	ohne Lehr- amtsoption <i>n</i> = 30	
Zeitdruck	1.79, .52	1.42, .51	$F(1, 89) = 10.15, p = .002, \eta^2 = .10$
Lernstrategien	2.08, .52	2.34, .38	$F(1, 89) = 5.94, p = .017, \eta^2 = .06$
Elaborations- muster	1.74, .48	1.51, .50	$F(1, 89) = 4.23, p = .043, \eta^2 = .05$

Ein weiteres Merkmal für subgruppenspezifische Varianzen stellt die bisher absolvierte Dauer im Vorbereitungsdienst dar (Tabelle 41). Hier zeigen sich Unterschiede zwischen den angehenden Lehrkräften, die sich noch in der Einführungsphase befinden und bisher noch nicht eigenständig unterrichten und denen, die sich bereits in der Phase des eigenständigen Unterrichts befinden (bedarfsdeckender Unterricht). Darauf bezogen sind auch ressourcenbezogene Differenzen erkennbar. Mit zunehmender Dauer des Vorbereitungsdienstes sammeln die angehenden Lehrkräfte nicht nur signifikant mehr Erfahrungen zum Gestalten von Lehr-Lern-Materialien, sondern auch ihre Selbstwirksamkeitserwartung steigt messbar an.

Tabelle 41: Signifikante Varianzanalysen zwischen angehenden Lehrkräften mit unterschiedlicher Dauer des bisher absolvierten bedarfsdeckenden Unterrichts

Variable	Dauer [<i>M</i> , <i>SD</i>]		Ergebnis der Varianzanalyse
	0 Monate <i>n</i> = 29	1-12 Monate <i>n</i> = 46	
Erfahrung mit LLM	1.68, .52	2.97, .68	$F(1, 89) = 6.71, p < .001, \eta^2 = .24$
SWE an sich selbst	2.91, .47	3.77, .57	$F(1, 89) = 5.89, p < .001, \eta^2 = .22$

Ein letztes relevantes Merkmal ist die Dauer einer Tätigkeit als Vertretungslehrkraft vor Antritt des Vorbereitungsdienstes. Auch hier zeigen sich Unterschiede

zwischen denjenigen, die keine entsprechende Unterrichtserfahrung gesammelt haben, und jenen, die bis zu 24 Monate als Vertretungslehrkraft tätig waren (Tabelle 42). Hier zeigt sich, dass mit zunehmender Unterrichtserfahrung die wahrgenommenen Belastungen und der empfundene Zeitdruck bei den angehenden Lehrkräften abnehmen.

Tabelle 42: Signifikante Varianzanalysen zwischen angehenden Lehrkräften mit und ohne einer Vertretungsstelle vor dem Vorbereitungsdienst

Variable	Dauer [<i>M</i> , <i>SD</i>]		Ergebnis der Varianzanalyse
	0 Monate <i>n</i> = 54	3-24 Monate <i>n</i> = 36	
Zeitdruck	1.80, .53	1.55, .63	$F(1, 89) = 3.032, p = .034, \eta^2 = .10$
Berufliche Belastung	1.06, .45	1.43, .41	$F(1, 89) = 5.25, p = .002, \eta^2 = .16$

Anmerkung. Dauer = Dauer der Übernahme einer Vertretungsstelle.

5.6.2 Korrelationen in der Gesamtgruppe zwischen allen Variablen

Die dargestellten Korrelationen wurden berechnet, um einerseits weitere Merkmale innerhalb der Gesamtgruppe aufzuklären und andererseits Hinweise auf die Konstruktvalidität, insbesondere auf die konvergente und diskriminante Validität, der erfassten Variablen zu untersuchen.

Im Folgenden werden die Zusammenhänge innerhalb der persönlichen Ressourcen (vgl. Kapitel 5.1.2) betrachtet (Tabelle 43). Auffällig ist, dass der wahrgenommene Zeitdruck die meisten signifikanten Korrelationen aufweist und einen starken Einfluss auf die weiteren Belastungsskalen hat. Je höher der Zeitdruck ist, desto höher ist die berufliche Belastung und desto weniger Elaborationsmuster und soziale Ressourcen werden genutzt. Als weiterer Faktor können die sozialen Ressourcen genannt werden, d. h. ob beispielsweise Mitreferendar:innen, Mentor:innen, Kolleg:innen und Fachleitungen bei Problemen oder Hindernissen um Rat gefragt werden. Der Einsatz sozialer Ressourcen korreliert nicht nur mit dem Zeitdruck, sondern auch mit dem Belastungserleben, dem Einsatz von Lernstrategien und der wahrgenommenen Wichtigkeit der Wissenschaft für die

Praxis. Insgesamt können daher der Zeitdruck und der Einsatz persönlicher Ressourcen als Maßstab gesehen werden, um die Belastung und Fähigkeit der Bewältigung angehender Lehrkräfte zu deuten. Auffällig ist, dass die Erfahrungen im Designen von Lehr-Lern-Materialien stark mit der Selbstwirksamkeit an den eigenen Chemieunterricht korrelieren. Dies kann so interpretiert werden, dass mit zunehmender Erfahrung auch die Selbstwirksamkeitserwartung steigt und sich die angehenden Lehrkräfte fähiger fühlen, lernwirksamen Unterricht durchzuführen.

Tabelle 43: Spearman-Korrelationen innerhalb der persönlichen Ressourcen

Variable	2.	3.	4.	5.	6.	7.	8.	9.
<i>Belastungen und Bewältigungsstrategien</i>								
1. Zeitdruck	.53**	.03	-.31**	-.33**	-.05	-.08	.14	-.31**
2. Berufl. Belastung		-.21	-.43**	-.26*	-.15	.18	-.02	-.16
3. Lernstrategien			.17	.38**	.33**	.15	.07	.14
4. Elaborationsmuster				.21	.17	-.20	.06	.10
5. Soz. Ressourcen					.23*	-.13	.15	-.00
<i>Professionelles Selbstverständnis</i>								
6. Wichtigkeit Wiss.						.08	.05	-.03
7. SWE an sich selbst							.10	.03
8. SWE an eigenen CU								.70**
<i>9. Erfahrungen LLM</i>								

Anmerkung. CU = Chemieunterricht, LLM = Lehr-Lern-Material, SWE = Selbstwirksamkeitserwartung. Hervorgehoben bzw. grau hinterlegt sind signifikante mittlere bis hohe Korrelationen.

Als nächstes werden die Korrelationen innerhalb der Wissenstests untersucht (Tabelle 44). Die in der Tabelle blau markierten Felder zeigen, in welchem Ausmaß die Variablen vom Prä- zum Posttest mit sich selbst korrelieren. Wie auch in früheren Studien zu beobachten (z. B. Dollny, 2011), lassen sich signifikante und mittelstarke bis starke Korrelationen feststellen, was auf eine hohe Konstruktvalidität hinweist (Bühner, 2021; Kapitel 6.1.1).

Tabelle 44: Pearson-Korrelationen zwischen den untersuchten Wissensbereichen im Prä- und Posttest

Variable	<i>Prätest</i>			<i>Posttest</i>			
	Allg. FDW	Rep. FW	Rep. FDW	Allg. FW	Allg. FDW	Rep. FW	Rep. FDW
<i>Prätest</i>							
Allg. FW	.20	.38**	.30**	.63**	.29**	.38**	.17
Allg. FDW		.24*	.13	.08	.38**	.13	.22*
Rep. FW			.33**	.30**	.30**	.60**	.33**
Rep. FDW				.22*	.21*	.03	.61**
<i>Posttest</i>							
Allg. FW					.34**	.37**	.22*
Allg. FDW						.36**	.23*
Rep. FW							.07
Rep. FDW							

Anmerkung. FW = Fachwissen, FDW = Fachdidaktisches Wissen. Die grün hinterlegten Ergebnisse beziehen sich auf Korrelationen zwischen den fachlichen und fachdidaktischen Variablen zum gleichen Messzeitpunkt. Gelb hinterlegte Ergebnisse stellen Korrelationen zwischen den allgemeineren und den repräsentationsbezogenen Variablen zum selben Messzeitpunkt dar. Blau markierte Ergebnisse zeigen den Prä-Post-Vergleich der gleichen Variablen.

Dieser Befund wird ebenso durch die Korrelationen zwischen den Variablen gestützt: Es lassen sich mehrere signifikante Korrelationen erkennen, bei denen der allgemeine Wissensbereich mit dem repräsentationsbezogenen Pendant zusammenhängt (Beispiel: Allgemeines und repräsentationsbezogenes Fachwissen im Prä- und auch Posttest). Auffällig ist, dass die fachlichen und fachdidaktischen Variablenpaare weniger stark bis gar nicht miteinander korrelieren (Beispiel: Repräsentationsbezogenes Fachwissen und fachdidaktisches Wissen im Prätest im Vergleich zum Posttest). Dies kann beispielsweise daran liegen, dass während der Intervention der Fokus auf dem fachdidaktischen Wissensbereich lag und nur dieser signifikant anstieg.

Nachdem die persönlichen Ressourcen und Wissensbereiche jeweils einzeln und innerhalb ihrer Domäne analysiert worden sind, folgen nun Korrelationen, die Hinweise darauf geben können, wie die beiden Domänen sich gegenseitig

beeinflussen. Dargestellt wird der Einfluss der persönlichen Ressourcen jeweils auf die allgemeinen (Tabelle 45) und repräsentationsbezogenen Wissensbereiche (Tabelle 46). Werden beide Tabellen betrachtet, lässt sich zunächst feststellen, dass viele Variablen aus den Ressourcen mit den Fachwissenstests korrelieren, während im fachdidaktischen Teil nur jeweils eine schwache Korrelation zu erkennen ist. Am stärksten korreliert hierbei das allgemeine Fachwissen sowohl im Prä- und Posttest mit den Ressourcen.

Tabelle 45: Spearman-Korrelationen zwischen den persönlichen Ressourcen und allgemeinen Wissensbereichen

Variable	Allg. Fachwissen		Allg. fachdid. Wissen	
	prä	post	prä	post
Zeitdruck	.37**	.26*	.02	.14
Beruf. Belastung	-.26*	-.20	-.19	-.11
Lernstrategien	-.01	.11	-.03	-.08
Elaborationsmuster	.17	.26*	.12	.11
Soz. Ressourcen	.22*	.23*	.10	.16
Wichtigkeit von Wiss.	.05	.18	.11	-.05
SWE an sich selbst	-.20	-.04	-.02	.09
SWE an eigenen CU	.32**	.29**	.17	.22*
Erfahrungen LLM	.31**	.25*	.02	.08

Anmerkung. CU = Chemieunterricht, LLM = Lehr-Lern-Material, SWE = Selbstwirksamkeitserwartung. Die grau hinterlegten Ergebnisse zeigen signifikante mittelstarke bis sehr starke Korrelationen.

Tabelle 46: Spearman-Korrelationen zwischen den persönlichen Ressourcen und den repräsentationsbezogenen Wissensbereichen

Variable	Rep. Fachwissen		Rep. fachdid. Wissen	
	prä	post	prä	post
Zeitdruck	.19	.11	.10	.19
Beruf. Belastung	-.24*	-.17	-.18	-.29**
Lernstrategien	.13	.15	-.05	-.02
Elaborationsmuster	.00	.00	.18	.17
Soz. Ressourcen	.30**	.24*	.03	.08
Wichtigkeit von Wiss.	.03	.07	.08	.19
SWE an sich selbst	-.12	-.09	-.07	.06
SWE an eigenen CU	.13	.28**	.09	.05
Erfahrungen LLM	.12	.26*	.13	.04

Anmerkung. CU = Chemieunterricht, LLM = Lehr-Lern-Material, SWE = Selbstwirksamkeitserwartung. Die grau hinterlegten Ergebnisse zeigen signifikante mittelstarke bis sehr starke Korrelationen.

Tabelle 47: Pearson-Korrelationen zwischen den Wissensbereichen sowie persönlichen Ressourcen und den Qualitäten der selbstdesignten Lehr-Lern-Materialien

Variable	Qualität von LLM		
	Nr. 1	Nr. 2	Nr. 3
Allg. Fachwissen (prä)	.34*	.09	.12
Allg. Fachwissen (post)	.35**	.23*	.18
Rep. Fachwissen (prä)	.30**	.02	.14
Rep. Fachwissen (post)	.22*	.05	.20
Rep. Überzeugungen (prä)	.31*	.35*	.33*
Rep. Überzeugungen (post)	.14	.18	.08
Wichtigkeit von Wissenschaft	.31**	.30*	.34*

Anmerkung. CU = Chemieunterricht, LLM = Lehr-Lern-Material, SWE = Selbstwirksamkeitserwartung. Die grau hinterlegten Ergebnisse zeigen signifikante mittelstarke bis sehr starke Korrelationen.

Als weitere abhängige Variable wurde im Hinblick auf die Qualität der Lehr-Lern-Materialien untersucht, inwiefern sie mit den erhobenen Wissensbereichen und Überzeugungen zusammenhängen (Tabelle 47). Die Qualität der Lern-Materialien weist zwischen den Sitzungen eine mittelstarke Korrelation auf (Nr. 1/2: $r = .45^{**}$, Nr. 2/3: $r = .41^{**}$, Nr. 1/3: $r = .40^{**}$), was darauf hinweist, dass sich die Qualitäten im Verlauf der Intervention nicht signifikant verbessern. Auffällig ist jedoch, dass die Qualität in der ersten Sitzung schwach bis mittelstark mit dem allgemeinen sowie repräsentationsbezogenen Fachwissen zusammenhängt. Dieser Effekt nimmt in der zweiten Sitzung ab und ist in der dritten Sitzung gar nicht mehr nachweisbar. Stattdessen bleiben die Korrelationen zwischen den repräsentationsbezogenen Überzeugungen und der empfundenen Wichtigkeit der Wissenschaft für die Schulpraxis hinweg konstant auf mittlerer Stärke.

Eine weitere abhängige Variable aus den Forschungsfragen, die es nach Zusammenhängen zu untersuchen gilt, sind die repräsentationsbezogenen Überzeugungen. Die Werte der angehenden Lehrkräfte aus dem Prätest korrelieren mit denen aus dem Posttest um $r = .64^{**}$ miteinander, was erneut auf eine hohe Konstruktvalidität hindeutet. Auffällig bei den Überzeugungen ist, dass sie nur mit einem bestimmten Teil der persönlichen Ressourcen signifikant korrelieren (Tabelle 48). Hierbei handelt es sich um alle Skalen aus dem professionellen Selbstverständnis. Die Überzeugungen der angehenden Lehrkräfte sind daher unabhängig von Belastungen und Bewältigungsstrategien und hängen eher damit zusammen, wie wichtig die angehenden Lehrkräfte wissenschaftliche Tätigkeiten, Befunde sowie Notwendigkeiten persönlicher Weiterentwicklungen betrachten und – viel ausschlaggebender – wie selbstwirksam sie ihre eigenen Tätigkeiten und Entwicklungen einschätzen.

Tabelle 48: Spearman-Korrelationen zwischen den repräsentationsbezogenen Überzeugungen und dem professionellen Selbstverständnis

Variable	Rep. Überzeugungen	
	prä	post
Wichtigkeit von Wissenschaft	.18	.25*
SWE an sich selbst	.42**	.32**
SWE an eigenen CU	.22*	.10

Anmerkung. CU = Chemieunterricht, SWE = Selbstwirksamkeitserwartung. Die grau hinterlegten Ergebnisse zeigen signifikante mittelstarke bis sehr starke Korrelationen.

Der wahrgenommene Cognitive Load und die Anstrengungsbereitschaft während des Arbeitens mit dem Bewertungsbogen wurden nach Korrelationen untersucht, um Hinweise auf eine mediierende Funktion zu finden. Die gemessenen drei Werte zum Cognitive Load in jeder Interventionssitzung korrelieren in einem hohen Grad miteinander (Nr. 1/2: $r = .66^{**}$, Nr. 2/3: $r = .74^{**}$, Nr. 1/3: $r = .58^{**}$). Gleiches gilt für die Anstrengungsbereitschaft (Nr. 1/2: $r = .53^{**}$, Nr. 2/3: $r = .42^{**}$, Nr. 1/3: $r = .41^{**}$). Beim Cognitive Load gibt es keine genauen Muster bei den Korrelationen (Tabelle 49). Signifikante Werte zeigen sich bei drei Skalen zu Ressourcen und dem allgemeinen fachdidaktischen Wissen. Auffällig ist hier, dass die Korrelationen mit den ersten drei Werten mit fortschreitender Interventionszeit abnehmen, wogegen die Korrelation mit dem fachdidaktischen Wissen sich verstärkt und nahezu verdoppelt. Bezogen auf die Anstrengungsbereitschaft wurden keine relevanten Korrelationen festgestellt. Mediierende Funktionen der beiden Variablen wurden in weiteren Analysen untersucht und konnten nicht festgestellt werden.

Als letzte Variablen werden die Abschlussnoten der Hochschulreife sowie des Hochschulabschlusses untersucht. Die Zahlenwerte der Noten wurden umgepolt, sodass höhere Werte einer besseren Leistung entsprechen. Eine positive Korrelation kann somit als Verbesserung beider Variablen in eine günstige Richtung interpretiert werden. Die Abiturnote korreliert mit folgenden Variablen: Hochschulabschlussnote ($r = .33^{**}$), Selbstwirksamkeitserwartung an sich selbst ($r = -.23^{*}$), Erfahrungen im Designen von Lehr-Lern-Materialien ($r =$

.22*), allgemeines fachdidaktisches Wissen ($r_{prä} = .31^{**}$, $r_{post} = .22^*$) und repräsentationsbezogenes Fachwissen ($r_{prä} = .31^{**}$, $r_{post} = .31^{**}$). Die Hochschulabschlussnote korreliert wiederum mit folgenden Variablen: Erfahrungen im Designen von Lehr-Lern-Materialien ($r = -.22$), Vorhandensein von Elaborationsmustern ($r = -.24^*$), Selbstwirksamkeitserwartung an sich selbst ($r = -.23^*$), repräsentationsbezogenen Überzeugungen ($r_{prä} = -.34^{**}$, $r_{post} = -.12$), allgemeines fachdidaktisches Wissen ($r_{prä} = -.18$, $r_{post} = -.36^{**}$), repräsentationsbezogenes Fachwissen ($r_{prä} = -.11$, $r_{post} = -.25^*$) und Qualität der Lehr-Lern-Materialien ($r_{Nr. 1} = -.21^*$, $r_{Nr. 2} = -.15$, $r_{Nr. 3} = -.23^*$).

Tabelle 49: Pearson-Korrelationen zwischen dem Cognitive Load und relevanten Variablen

Variable	Cognitive Load bei LLM		
	Nr. 1	Nr. 2	Nr. 3
Wichtigkeit von Wissenschaft	-.30*	-.29*	-.22
SWE an sich selbst	-.38**	-.39**	-.30*
Erfahrungen LLM	-.30*	-.25	-.15
Allg. fachdid. Wissen	-.28*	-.46**	-.49**

Anmerkung. LLM = Lehr-Lern-Material, SWE = Selbstwirksamkeitserwartung. Die grau hinterlegten Ergebnisse zeigen signifikante mittelstarke bis sehr starke Korrelationen.

5.6.3 Gruppenspezifische Median-Splits im Fachwissen und fachdidaktischen Wissen

Um die Korrelationen und daraus resultierende Effekte differenzierter zu untersuchen, wurden Median-Splits durchgeführt und anschließend Varianzanalysen in Bezug auf das repräsentationsbezogene Fachwissen und fachdidaktische Wissen berechnet. Ziel der Analyse war es zu untersuchen, inwiefern die beiden Wissensarten von dem Vorwissen abhängen und, inwiefern Gruppeneffekte erkennbar sind. Die Median-Splits wurden jeweils für alle vier Wissensbereiche durchgeführt. Anschließend wurde untersucht, ob das allgemeine und repräsentationsbezogene Wissen im Prätest das repräsentationsbezogene Wissen im Posttest beeinflusst. Die Ergebnisse werden anhand von Streudiagrammen visualisiert (Abbildung 36-Abbildung 39).

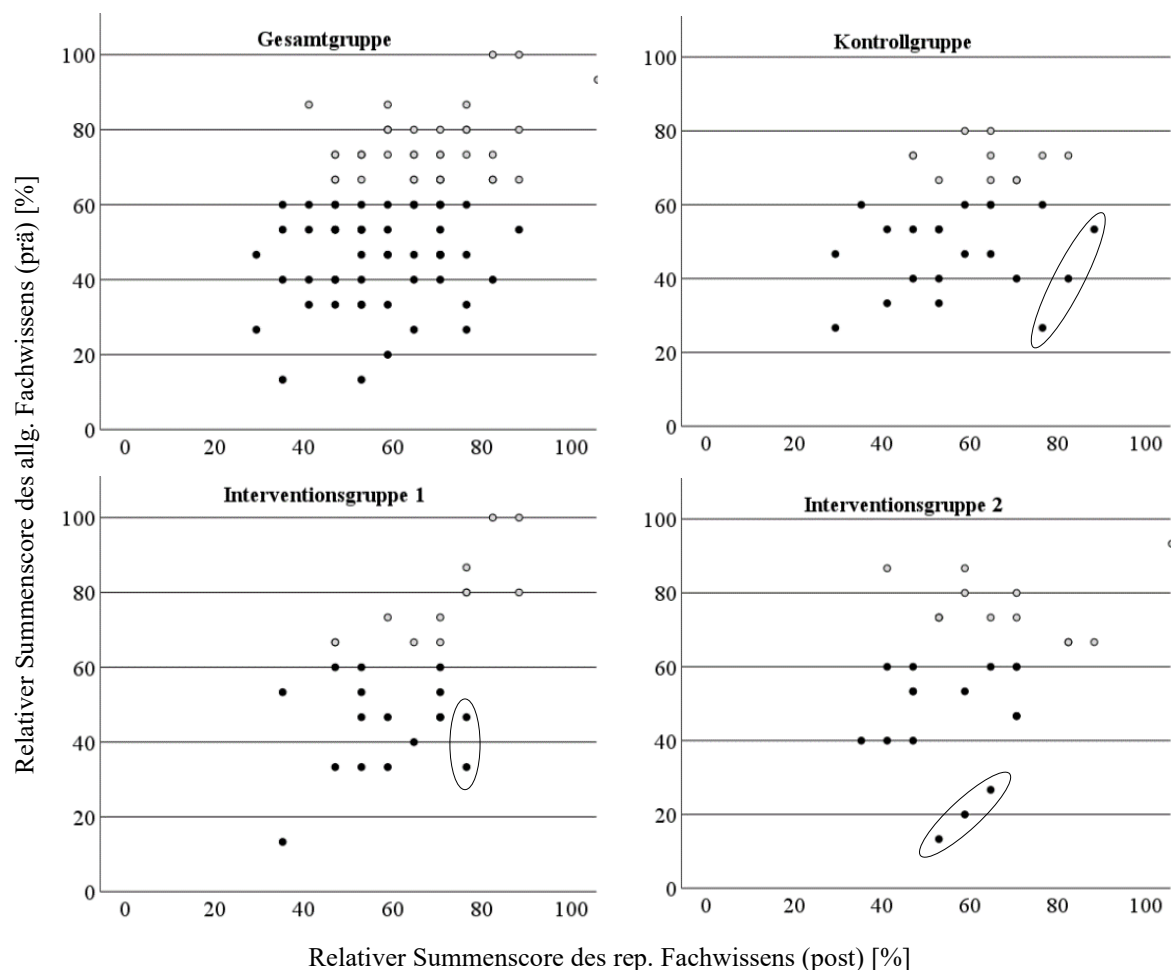


Abbildung 36: Streuung zwischen dem allgemeinen Fachwissen im Prätest und repräsentationsbezogenen Fachwissen im Posttest

In Abbildung 36 wird der Einfluss des allgemeinen Fachwissens im Prätest auf das repräsentationsbezogene Fachwissen im Posttest dargestellt. Die bereits berechnete moderate Korrelation von $r = .38, p < .01$ ist mit einer eher ovalen, flächigen Streuung der Daten in der Gesamtgruppe erkennbar. Werden die Daten nach den Untersuchungsgruppen aufgeteilt, so ist in der Kontrollgruppe ($r = -.23, p = .20$) und Interventionsgruppe 2 ($r = .32, p = .09$) eine ähnliche Streuung wie in der Gesamtgruppe erkennbar. Dagegen ist eine starke Korrelation zwischen den beiden Variablen in der Streuung der Interventionsgruppe 1 deutlich zu sehen ($r = .68, p < .001$). Eine Varianzanalyse zwischen den Low und High Performenden der Interventionsgruppe 1 bezüglich des allgemeinen Fachwissens indizieren einen Unterschied in der Performanz im repräsentationsbezogenen Fachwissen ($U = 53.50, p = .30, \eta^2 = .56$). Aufgrund der geringeren Substichprobengrößen wurden Varianzanalysen für nicht-parametrische Tests bzw. Mann-Whitney-U-Tests durchgeführt. Auffällig ist in allen drei Untersuchungsgruppen eine Anhäufung von Fällen, bei denen Low Performende des allgemeinen Fachwissens teilweise um das Dreifache bessere Leistungen im Posttest des repräsentationsbezogenen Fachwissens erzielen (Abbildung 36, eingekreiste Fälle in KG, IG1 und IG2).

Die Abhängigkeit des repräsentationsbezogenen Fachwissens im Posttest vom Prätest ist in der Abbildung 37 dargestellt. Hier ist für die Gesamtgruppe eine starke Korrelation von $r = .60, p < .01$ vorhanden. Diese Korrelation zeigt sich ebenso durch die Streuung in der Interventionsgruppe 1 ($r = .68, p < .001$). In der Kontrollgruppe ($r = .68, p < .001$) und Interventionsgruppe 2 ($r = .47, p = .011$) sind jedoch andere Muster erkennbar (Abbildung 37, eingekreiste Fälle in KG und IG2). Unabhängig von einem hohen oder niedrigen Vorwissen erreichen die angehenden Lehrkräfte hier ein gleich hohes Fachwissen. Beispielsweise gibt es in der Kontrollgruppe Fälle, bei denen eine Testperson im Vortest etwa 35 % erreicht, eine andere Testperson 70 % und beide im Posttest ca. 78 %. Sehr ähnliche Fälle sind auch in der Interventionsgruppe 2 vorhanden. Auskunft zu Ursachen bzw. entscheidenden unabhängigen Variablen, warum das Lernen in wenigen Einzelfällen unerwarteterweise doch stattfindet oder nicht, können Regressionsanalysen geben. Außerdem unterscheiden sich Low und High Performende der Interventionsgruppe 1 ($U = 29.00, p < .001, \eta^2 = .64$) und KG ($U = 23.00, p < .001, \eta^2 = .66$) in ihrer Performanz im Posttest.

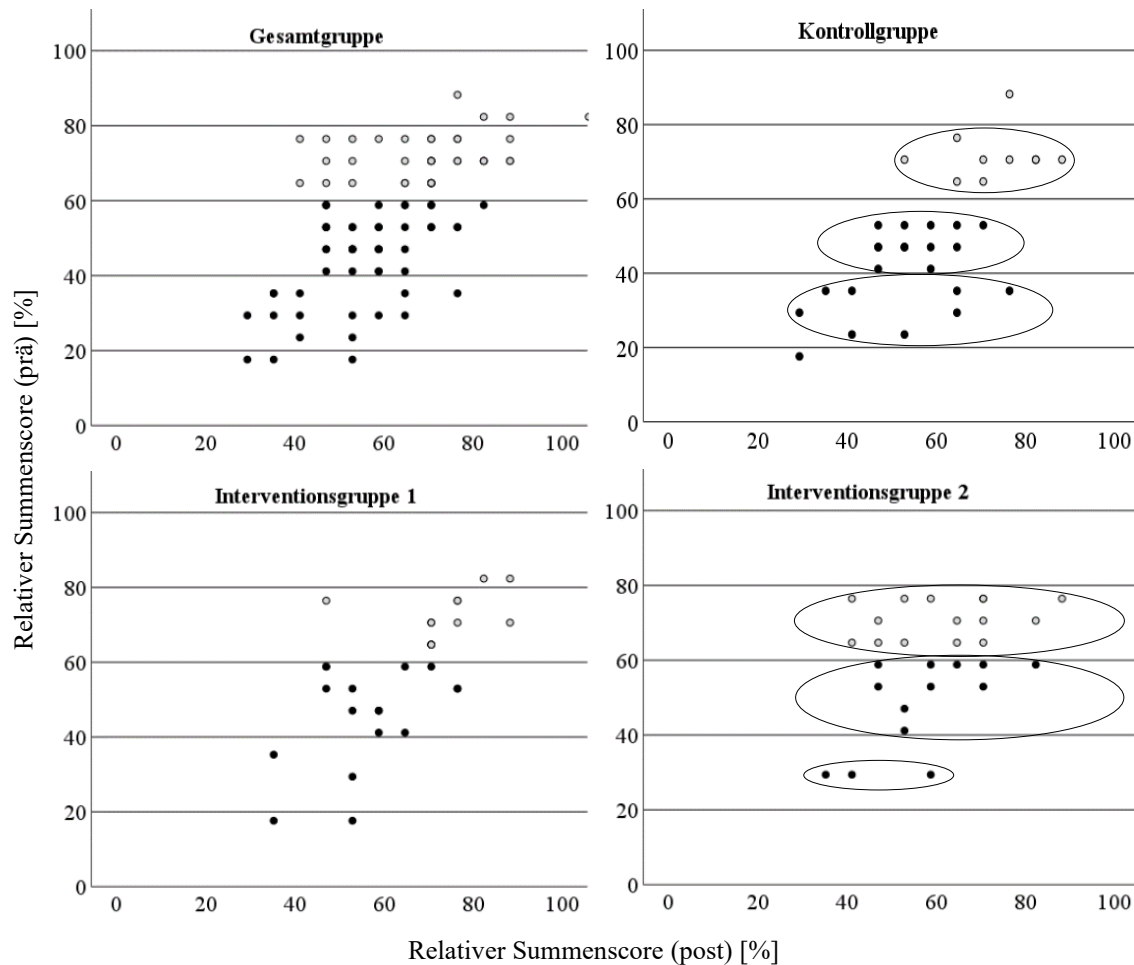


Abbildung 37: Streuung des repräsentationsbezogenen Fachwissens zwischen Prä- und Post-test

Bezogen auf das repräsentationsbezogene fachdidaktische Wissen zeigen sich andere Streuungsmuster. Zunächst wird die Abhängigkeit vom allgemeinen fachdidaktischen Wissen im Prätest betrachtet (Abbildung 38). Korrelationsanalysen zeigen einen schwachen Zusammenhang auf ($r = .22, p < .05$). Die Daten konzentrieren sich insgesamt in dem oberen linken Bereich des Streudiagramms und sind insgesamt so breit gestreut, dass keine sinnvolle Korrelationsgerade durch den Graphen gezogen werden kann. Eine hohe Leistung im allgemeinen fachdidaktischen Wissen geht daher nicht mit einem höheren repräsentationsbezogenen Fachwissen einher. Diese Unabhängigkeit lässt sich in der Kontrollgruppe ($r = .31, p = .087$) und Interventionsgruppe 2 ($r = .29, p = .126$) noch stärker erkennen (Abbildung 38, waagrecht eingekreiste Fälle in KG, IG1 und

IG2). Die zweite auffällige Ansammlung der Daten befindet sich in der Kontrollgruppe (senkrecht eingekreiste Fälle), in der zu erkennen ist, dass eine Gruppe von Testpersonen unabhängig von ihrem Vorwissen im Posttest vergleichsweise schlecht performen. Hier kann es beispielsweise sein, dass die Testpersonen unzufrieden mit der Intervention und dem wenig gewinnbringenden Dummy-Feedback sind und folglich schlecht performen. Die dritten und letzten auffälligen Ansammlungen sind drei Testpersonen aus der Interventionsgruppe 1 und eine Testperson aus der Interventionsgruppe 2, die im Vergleich zu einem geringen Vorwissen eher gut abschneiden. Es zeigt sich also erneut, dass es für den Wissenserwerb der angehenden Lehrkräfte noch andere Ursachen geben muss.

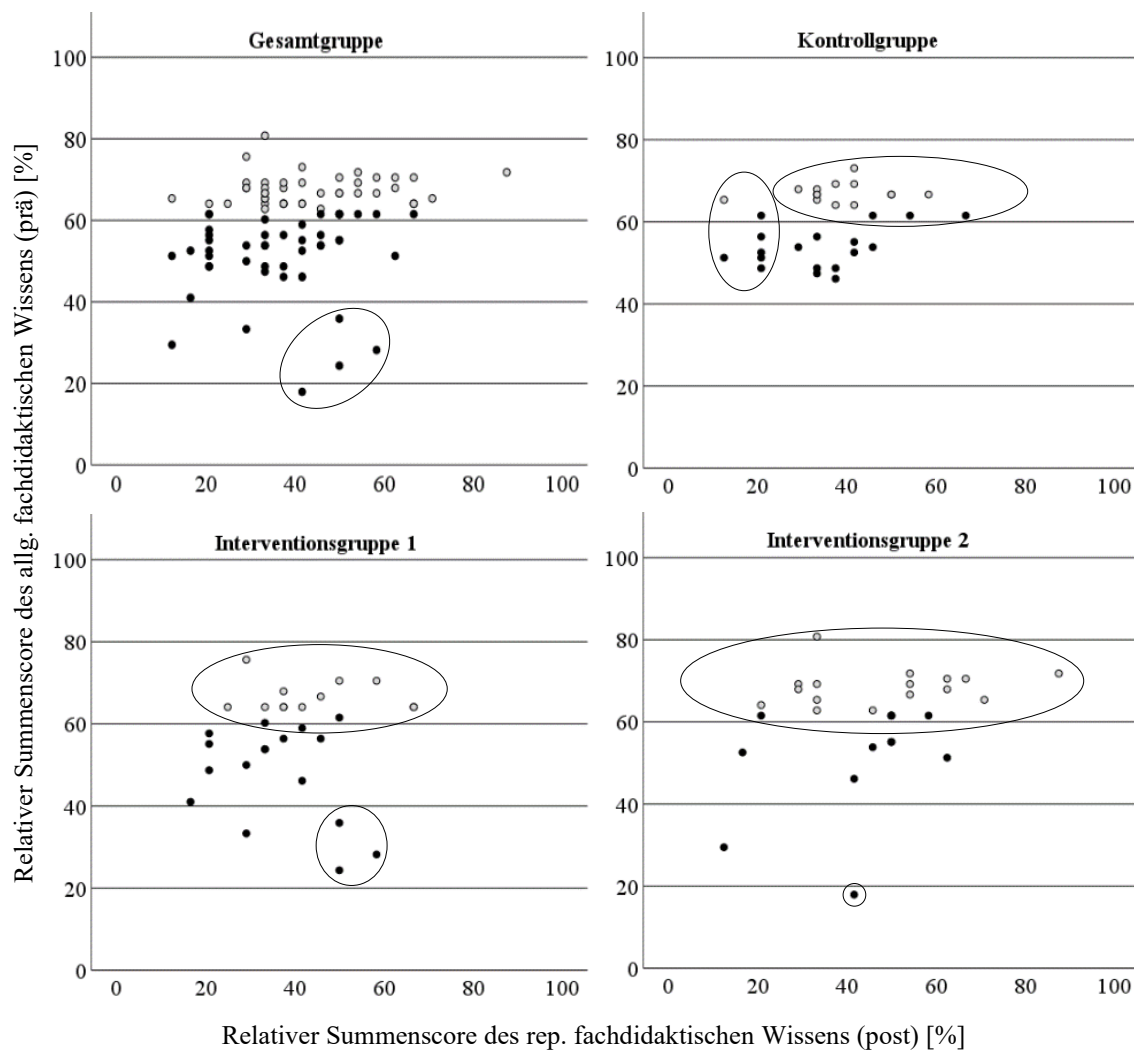


Abbildung 38: Streuung zwischen dem allgemeinen fachdidaktischen Wissen im Prätest und repräsentationsbezogenen fachdidaktischen Wissen im Posttest

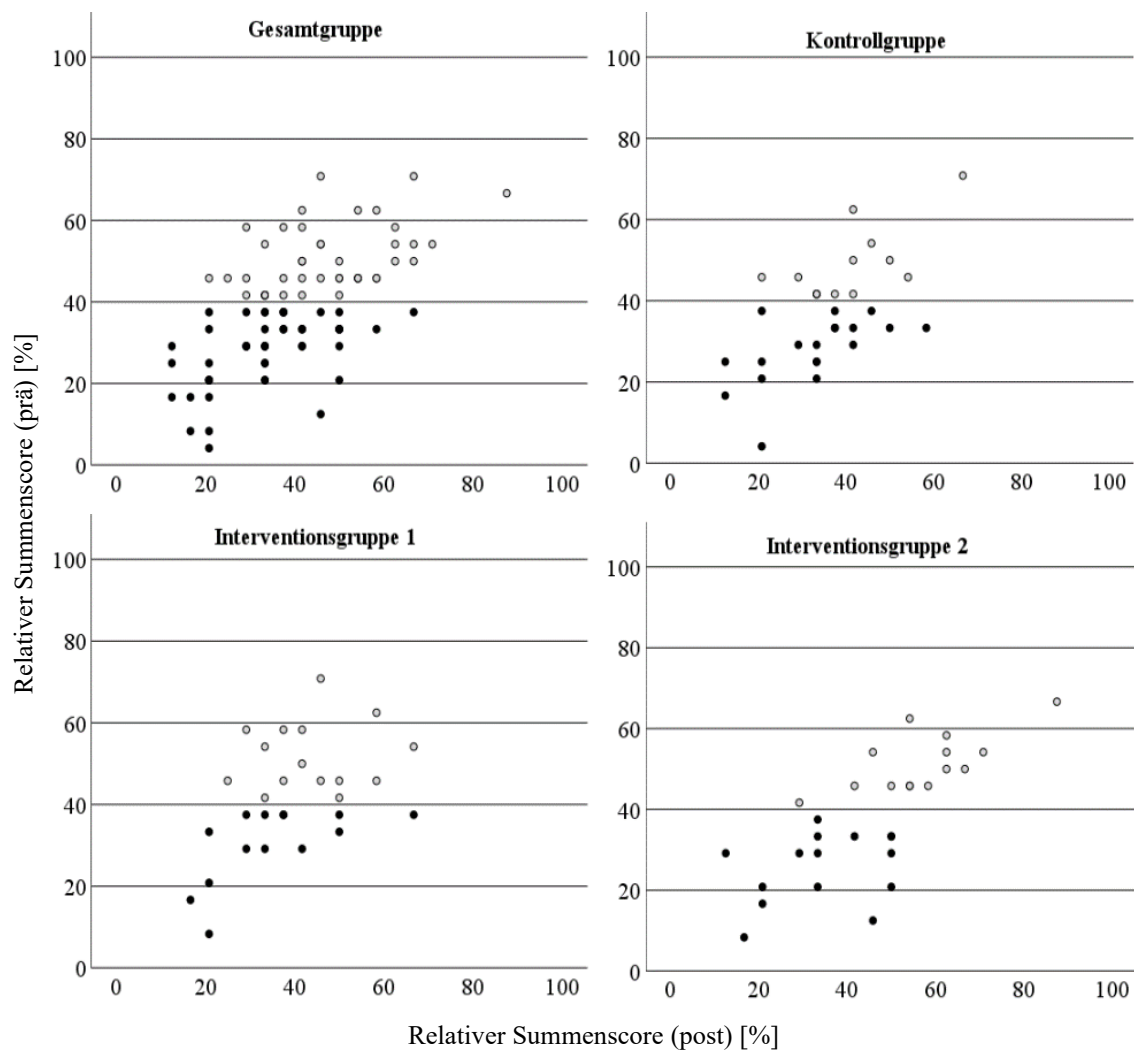


Abbildung 39: Streuung des repräsentationsbezogenen fachdidaktischen Wissens zwischen Prä- und Posttest

Die Abhängigkeit des repräsentationsbezogenen fachdidaktischen Wissens im Posttest vom jeweiligen Prätest in Abbildung 39 zeigt deutlich, dass sich die Daten im unteren linken Feld des Streudiagrammes häufen. Dies deutet darauf hin, dass die angehenden Lehrkräfte insgesamt ein geringes Vorwissen vor der Erhebung aufweisen und dass dieses nur begrenzt verbessert wurde. Diese geringe Veränderung zwischen Prä- und Posttest wird durch eine hohe Korrelation der beiden Variablen gestützt ($r = .61, p < .01$).

Eine sichtbare Verschiebung der Werte in Richtung des oberen rechten Diagrammbereichs lässt sich nur in der Interventionsgruppe 2 erkennen ($r = .75, p < .001$), wo Varianzanalysen ohnehin einen signifikanten Lernzuwachs indiziert

haben. Daher sind auch Unterschiede in der Performanz der Interventionsgruppe 2 zwischen Low und High Performenden vorzufinden ($U = 22.50$, $p < .001$, $\eta^2 = .66$). Die Korrelation innerhalb der Interventionsgruppe 1 beträgt $r = .48$, $p = .008$ und Kontrollgruppe $r = .61$, $p < .001$.

5.6.4 Prädiktoren des Fachwissens und fachdidaktischen Wissens

Die abschließenden explorativen Regressionsanalysen dienen zur Prognose der Kontrollvariablen und abhängigen Variablen. Abgebildet werden die Regressionen als Pfadanalyse in einem Pfaddiagramm. Die genauen Stufen der Modelle sowie ihre Kennwerte sind im Anhang der Arbeit dokumentiert.

Zunächst werden das allgemeine Fachwissen und das fachdidaktische Wissen der Gesamtgruppe untersucht (Abbildung 40). Hierzu werden die Daten aus dem Prätest verwendet, weil dabei noch keine unterschiedlichen Interventionen durchlaufen wurden und es sich bei den genannten Variablen um Kontrollvariablen handelt, die während der Intervention stabil geblieben sind. Die Prädiktoren repräsentationsbezogenes Fachwissen, Zeitdruck und Selbstwirksamkeitserwartung sagen allgemeines Fachwissen voraus ($F(82, 3) = 11.30$, $p < .001$, $N = 90$) und allgemeines fachdidaktisches Wissen ($F(83, 2) = 10.77$, $p < .001$, $N = 90$).

Bei den Prädiktoren des repräsentationsbezogenen Fachwissens wird ebenso die Gesamtstichprobe analysiert. Dies gilt sowohl für den Prä- als auch Posttest, weil keine gruppenspezifischen Varianzen nach der Intervention feststellbar waren. Die Prädiktoren zum repräsentationsbezogenen Fachwissen ändern sich durch die Intervention (Abbildung 41). Im Prätest kann das repräsentationsbezogene Fachwissen mit dem allgemeinen Fachwissen, dem Einsatz sozialer Ressourcen sowie dem repräsentationsbezogenen fachdidaktischen Wissen vorhergesagt werden ($F(82, 3) = 8.57$, $p < .001$, $N = 90$). Dies ändert sich im Posttest, wo neben dem repräsentationsbezogenen fachdidaktischen Wissen auch die Selbstwirksamkeitserwartung und das Vorwissen selbst an repräsentationsbezogenem Fachwissen als Prädiktoren gelten ($F(82, 3) = 19.73$, $p < .001$, $N = 90$).

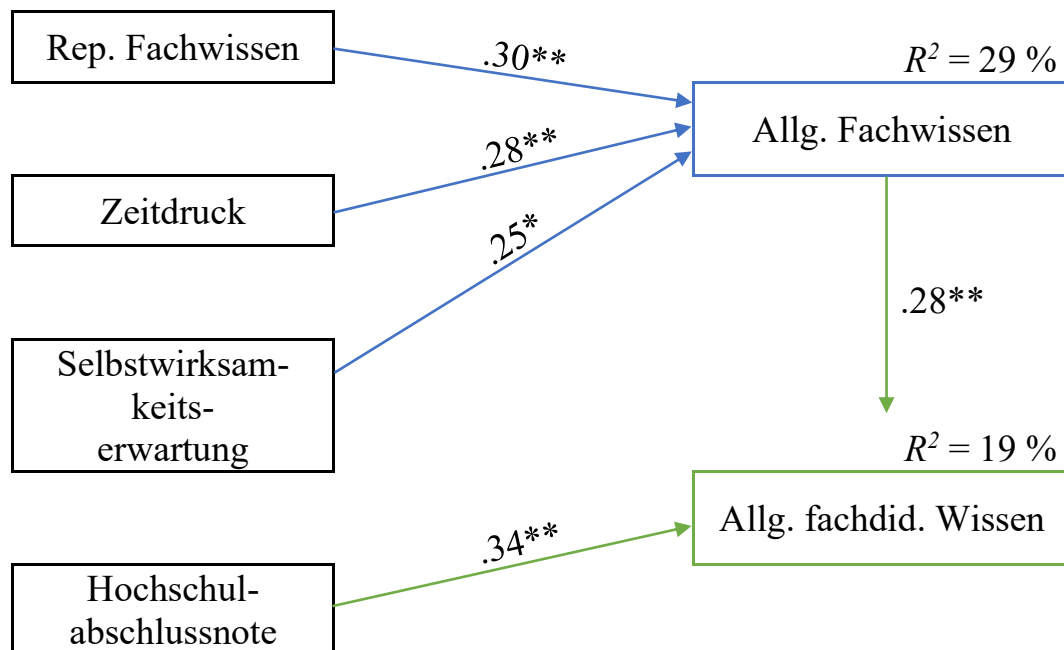


Abbildung 40: Prädiktoren des allgemeinen Fachwissens und fachdidaktischen Wissens mit zugehörigen β -Koeffizienten

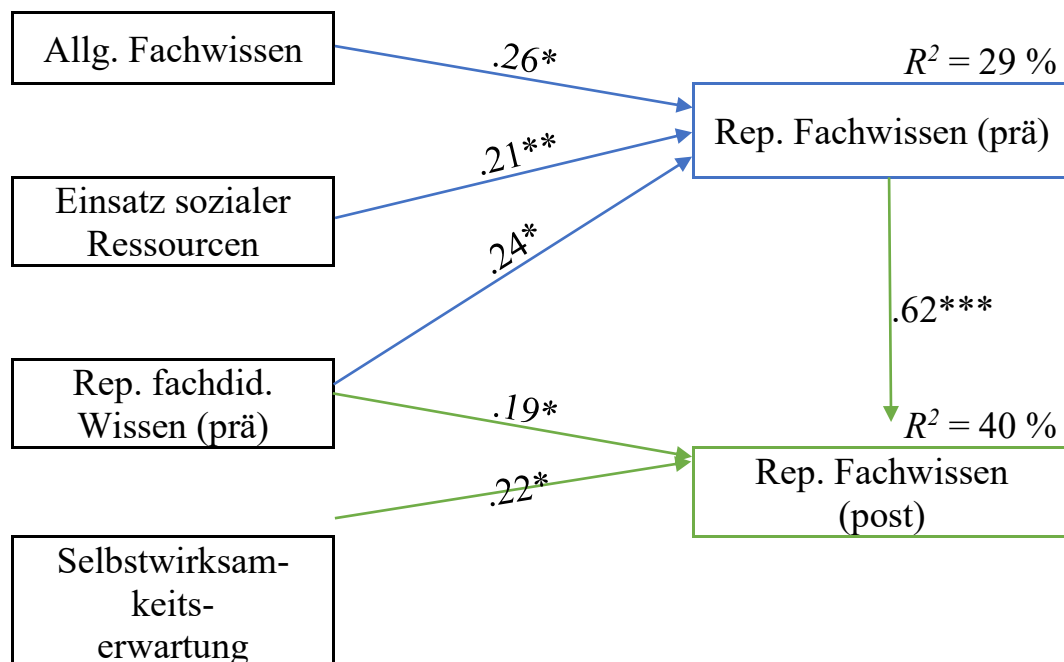


Abbildung 41: Prädiktoren des repräsentationsbezogenen Fachwissens im Prä- und Posttest mit zugehörigen β -Koeffizienten

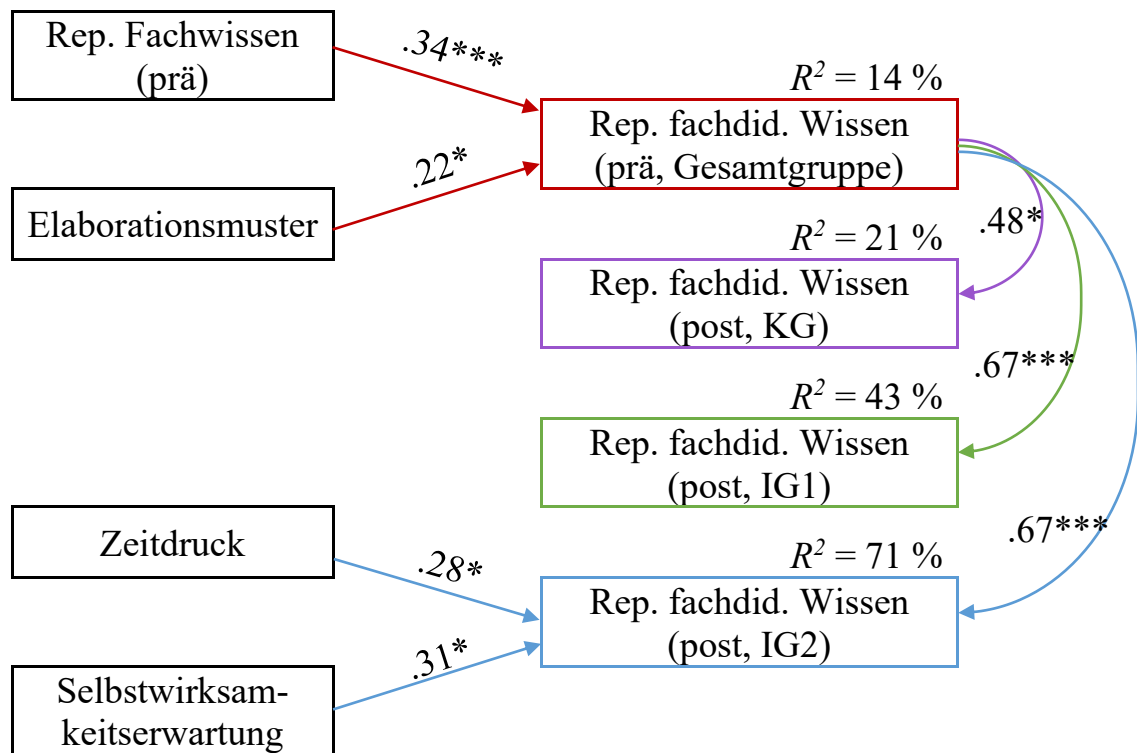


Abbildung 42: Prädiktoren des repräsentationsbezogenen fachdidaktischen Wissens mit zugehörigen β -Koeffizienten, im Posttest unterteilt nach den Untersuchungsgruppen

Eine Verschiebung der Prädiktoren ist auch beim repräsentationsbezogenen fachdidaktischen Wissen zu finden (Abbildung 42). Die Prädiktoren des Prätests wurden anhand der Gesamtstichprobe berechnet, weil zu diesem Messzeitpunkt bzw. vor der Intervention keine gruppenspezifischen Varianzen vorliegen. Das repräsentationsbezogene fachdidaktische Wissen lässt sich im Prätest durch das repräsentationsbezogene Fachwissen sowie den Einsatz von Elaborationsmustern vorhersagen ($F(83, 2) = 7.89, p < .001, N = 90$). Nach der Intervention spielen diese Kriterien keine Rolle mehr, sondern das repräsentationsbezogene fachdidaktische Wissen wird nun maßgeblich durch das Vorwissen bestimmt. Je mehr die angehenden Lehrkräfte also über das fachdidaktische Wissen bereits vorher verfügen, desto stärker nimmt es durch die Intervention zu. Dieser Prädiktor gilt im Posttest für alle drei Untersuchungsgruppen (KG: $F(27, 1) = 22.12, p < .001, n = 31$; IG1: $F(27, 1) = 8.24, p = .008, n = 29$). Bei der Inter-

ventionsgruppe 2 kommen aber noch zwei weitere Prädiktoren hinzu: der wahrgenommene Zeitdruck im Vorbereitungsdienst und die Selbstwirksamkeitserwartung ($F(24, 3) = 22.71, p < .001, n = 29$).

Die Prädiktoren für die repräsentationsbezogenen Überzeugungen werden mit den Daten aus dem Prätest aus dem Gesamtdatensatz berechnet, weil zum einen sich die Überzeugungen durch die Intervention nicht verändern und zum anderen es keine gruppenspezifischen Unterschiede gibt. Als Prädiktoren der Überzeugungen wurden die Selbstwirksamkeitserwartung ($\beta = .34, p < .001$) und die Hochschulabschlussnote ($\beta = .27, p < .01$) identifiziert ($F(83, 2) = 12.23, p < .001, R^2 = 23 \%, n = 90$).

Als letzte Untersuchungsvariable wurden Prädiktoren der Qualität der von den angehenden Lehrkräften designten Lehr-Lern-Materialien analysiert. Da sich die Qualität im Verlauf der Intervention nicht verbessert und auch keine Gruppeneffekte festgestellt wurden, erfolgten die Analysen erneut auf Basis der Gesamtstichprobe sowie des ersten Messzeitpunktes bzw. der Lehr-Lern-Materialien aus der ersten Sitzung. Die Regressionsanalysen zeigen, dass die Qualität der Lehr-Lern-Materialien maßgeblich durch das allgemeine Fachwissen der angehenden Lehrkräfte (prä, $\beta = .32, p < .01$) sowie durch ihre Einschätzung der Wichtigkeit der Wissenschaft für die Schulpraxis ($\beta = .28, p < .01$) prognostiziert werden kann ($F(83, 2) = 9.59, p < .001, R^2 = 17 \%, n = 90$).

6 DISKUSSION UND INTERPRETATION DER ERGEBNISSE

Im Rahmen des DFG-geförderten Forschungsprojekts *FiRe₂ – Feedback im Referendariat zu externen Repräsentationen* wurde der Forschungsfrage nachgegangen, inwieweit angehende Lehrkräfte des Unterrichtsfaches Chemie von internem und/oder externem Feedback profitieren können. Hierzu wurde eine Interventionsstudie im Prätest-Posttest-Kontrollgruppendesign durchgeführt sowie eine ausführliche Dokumentation individueller Bedingungen im Vorbereitungsdienst berichtet.

6.1 Methodische Reflexion

6.1.1 Güte der Testinstrumente

Die Auswahl, Adaption und Entwicklung der eingesetzten Testinstrumente nahm im Rahmen der vorliegenden Interventionsstudie eine zentrale Rolle ein. Ein wesentliches Ziel bestand darin, das repräsentationsbezogene Fachwissen und fachdidaktische Wissen sowie die repräsentationsbezogenen Überzeugungen möglichst valide und reliabel zu erfassen. Grundlage bildeten dabei die Testinstrumente aus dem ProwiN-Projekt (Dollny, 2011) und KiL-Projekt (Kleickmann et al., 2014; Taskin et al., 2015a) sowie weitere erprobte Erhebungsinstrumente anderer Fachbereiche (Abs et al., 2005; Nitz et al., 2011).

Die vorliegenden Tests wurden inhaltlich analysiert, angepasst und teilweise grundlegend modifiziert, um eine größtmögliche Passung zu den Zielsetzungen der Studie zu gewährleisten. In diesem Prozess lag ein besonderer Fokus auf der Sicherstellung der Inhaltsvalidität, sodass sowohl die Testinstrumente zum Fachwissen als auch das fachdidaktische Wissen möglichst für das Setting des Vorbereitungsdienstes im Fach Chemie ausgestaltet wurden. Dennoch zeigte sich, dass insbesondere bei den adaptierten bzw. neu entwickelten Items die Reliabilitätswerte je nach Untersuchungsvariable schwankten. Während im Bereich des allgemeinen Fachwissens und der Überzeugungen akzeptable bis gute interne Konsistenzen erreicht wurden ($\alpha > .7$; vgl. Blanz, 2021), gestaltete sich dies für das repräsentationsbezogene fachdidaktische Wissen aufgrund der Vielzahl an neuen oder stark modifizierten Items anspruchsvoller ($\alpha = .6-.7$). Hier

zeigt sich eine vergleichbare Bandbreite an Reliabilitätswerten wie auch in ähnlichen empirischen Studien zum Professionswissen von Lehrkräften (Dollny, 2011; Kleickmann et al., 2014).

Für die Items zur Erfassung der Überzeugungen wurde durch Überarbeitung und Erweiterung (z. B. formale Anpassungen, sprachliche Präzisierung, fachlicher Kontextualisierung) ein direkter und konkreter Bezug zum Chemieunterricht hergestellt. Nach einer Pilotierung und detaillierten Itemanalyse konnten die Werte für die interne Konsistenz in der Hauptstudie deutlich gesteigert werden. Dennoch traten in der Stichprobe Deckeneffekten auf, die Anlass dafür geben, repräsentationsbezogene Überzeugungen von Lehrkräften weiter zu erforschen. Die Testung der Qualität der designten Lehr-Lern-Materialien erfolgte mit einem in der Studie entwickelten Bewertungsbogen, der in mehreren Abschlussarbeiten empirisch und praktisch erprobt wurde (z. B. Mühlhaus, 2020; Tonyali, 2018). Die Ratingskala ließ sich nach statistischen Analysen sowohl bezüglich ihrer Benutzerfreundlichkeit als auch ihrer objektiven Anwendbarkeit als geeignet bewerten. Darüber hinaus wurde die Funktion der Skala als ein geeignetes Werkzeug zur Selbsteinschätzung durch die Ergebnisse der angehenden Lehrkräfte bestätigt.

Insgesamt erfüllen die eingesetzten Testinstrumente daher die Anforderungen an die Gütekriterien von empirischen und quantitativen Erhebungen (nach Döring & Bortz, 2023). Durch die Prozessschleife von Pilotierung, fortlaufender Analyse der Testkennwerte und Nachjustierung konnte die Qualität der Instrumente und Materialien kontinuierlich erhöht werden. Dennoch sollte als Ausblick betont werden, dass weitere Optimierungen, beispielsweise durch gezieltere Itemanalysen in größeren und heterogeneren Stichproben ein Potenzial zur weiteren Steigerung der Güte bieten.

6.1.2 Eignung des Studiendesigns und Umsetzung der Intervention

Das Studiendesign der vorliegenden Arbeit basiert auf einem experimentellen Prä-Post-Kontrollgruppendesign, das auf die Untersuchung der Bedingung und Wirkung des internen und externen Feedbacks im Vorbereitungsdienst ausgerichtet war. Die Randomisierung der Teilnehmenden in Kombination mit der Gruppenzuteilung nach Lehramtsform und Studienhintergrund stellen methodische Stärken dar. Auf diese Weise konnten mögliche systematische Verzerrungen aufgrund demografischer Daten und Ausgangssituationen minimiert werden

(nach Döring & Bortz, 2023). Auch die Varianzhomogenität der demografischen Daten und des Vorwissens unterstreicht die Annahme, dass die Untersuchungsgruppen die Intervention mit vergleichbaren Ausgangsbedingungen durchliefen.

Die Integration der Intervention in den Vorbereitungsdienst während der Corona-Pandemie kann einerseits als eine reale Ausbildungssituation und andererseits aber auch als eine weltweite Ausnahmesituation betrachtet werden. Die Integration im Vorbereitungsdienst ermöglichte dennoch eine hohe Authentizität und Praxisnähe. Durch die Wahl onlinebasierter und asynchroner Sitzungen mit Hilfe der Lernplattform Moodle konnte auf die pandemiebedingten Einschränkungen flexibel und schnell reagiert werden. Dies erlaubte nicht nur eine größere örtliche und zeitliche Unabhängigkeit für die angehenden Lehrkräfte, sondern erhöhte auch die Standardisierung der Durchführungsbedingungen (z. B. Bearbeitungszeit, Anleitung der Intervention). Dementsprechend konnte auch eine größere Stichprobengröße realisiert werden, als es in klassischen Präsenzstudien unter den damaligen Rahmenbedingungen möglich gewesen wäre (vgl. Nitz et al., 2011).

Jedoch sind gleichzeitig auch die deutlichen Herausforderungen des vorliegenden Vorgehens zu beachten. So brachte die asynchrone Bearbeitung einen erhöhten Kontrollverlust über die tatsächlichen Arbeitszeiten und -prozesse mit sich. Die Eigenverantwortlichkeit der Teilnehmenden führte möglicherweise zu Varianzen in der Motivation und in der Intensität der Bearbeitung. Methodisch wurde dem durch die Erhebung der wahrgenommenen Anstrengungsbereitschaft und des Cognitive Load entgegengewirkt. Die Angaben der angehenden Lehrkräfte könnten jedoch von sozialer Erwünschtheit beeinflusst worden sein. Außerdem bleibt festzuhalten, dass sich Schwierigkeiten, wie Dropout, unterschiedliche technische Voraussetzungen und individuelle Ablenkungen nicht kontrollieren ließen (Döring & Bortz, 2023).

Ein weiterer Aspekt ist die Gestaltung der Feedbackbedingungen. Die klare Trennung der Konzeptionen von internem Feedback, externem Feedback und Dummy-Feedback erlaubte eine differenzierte Analyse der Feedbackwirkungen. Der Ausschluss einer möglichen Interventionsgruppe, die nur mit externem Feedback und ohne internes Feedback arbeiten sollte, konnte mit dem ITFL-

Modell von Narciss (2018) weitreichend begründet werden, da externes Feedback zwangsläufig Prozesse des internen Feedbacks initiiert.

Die Intervention selbst wurde so aufgebaut, dass sie aus mehreren Zyklen aus Übungsphase, (Selbst-)Bewertung und Reflexion bestand. Dieses Vorgehen ermöglichte die systematische Entwicklung von professionellen Kompetenzen und zielte auf die kontinuierliche Verbesserung der Materialgestaltung ab. Die gewählte Aufgabenstellung, aus bereitgestellten Rohmaterialien praxisnahe Lehr-Lern-Materialien zu gestalten, förderte die Authentizität, steigerte jedoch zugleich die methodische und statistische Komplexität und Vergleichbarkeit der individuellen Arbeitsprodukte.

Abschließend ist die interne und externe Validität des vorliegenden Studiendesigns zu reflektieren. Die interne Validität bzw. Kohärenz innerhalb der Studie wurde durch verschiedene Maßnahmen gestärkt. Dazu zählen die Randomisierung der Proband:innen zu den Untersuchungsgruppen, die Varianzhomogenität in den Prätests, der systematische Einsatz valider und reliabler Testinstrumente sowie die Standardisierung der Intervention mittels digitaler Lernumgebung. Durch die parallele Erhebung und Kontrolle individueller Ausgangsdaten konnten potenzielle Störfaktoren (z. B. Vorwissen, Ressourcen, Ausbildungshintergrund) kontrolliert werden. Durch die möglichst transparente Dokumentation aller Abläufe und die umfassenden Kontrollmaßnahmen konnten die Einschränkungen zudem weitestgehend adressiert und auf ein akzeptables Maß begrenzt werden.

Die externe Validität, also die Generalisierbarkeit der Befunde auf andere Kontexte, ist im vorliegenden Setting jedoch eingeschränkt: Die Fokussierung auf das Fach Chemie, die Rekrutierung der Stichprobe aus einem Bundesland (NRW), der Vorbereitungsdienst und die Intervention unter Pandemiebedingungen sowie die Durchführung als rein digitale, asynchrone Intervention limitieren die Übertragbarkeit der Ergebnisse auf andere Bundesländer, Unterrichtsfächer und Bildungsphasen. Für weiterführende Generalisierungen wären daher vergleichende Studien in anderen Fachrichtungen und Rahmenbedingungen notwendig. Insgesamt ist somit die interne Validität der vorliegenden Studie als hoch, die externe Validität als begrenzt einzuschätzen.

6.2 Diskussion und Beantwortung der Forschungsfragen

Im Folgenden werden die zentralen Befunde der Studie im Hinblick auf die Forschungsfragen und Hypothesen diskutiert. Jeder Abschnitt beginnt mit einer knappen Zusammenfassung der empirischen Ergebnisse aus Kapitel 5 und ordnet diese anschließend im Kontext der bestehenden Forschung und Theorie ein. In Hinblick auf die einzelnen Forschungsfragen können daher folgende Erkenntnisse festgehalten werden:

6.2.1 *F_{1A}: Effekt von internem und/oder externem Feedback auf das repräsentationsbezogene Fachwissen und fachdidaktische Wissen von angehenden Lehrkräften*

Die Auswertungen und Analysen zeigen, dass repräsentationsbezogenes Fachwissen über alle Gruppen hinweg auf mittlerem Niveau stabil bleibt und kein signifikanter Wissenszuwachs durch die Intervention erzielt wird. Das repräsentationsbezogene fachdidaktische Wissen hingegen zeigt einen signifikanten Anstieg in der Interventionsgruppe 2, die mit ausschließlich internem Feedback arbeitete. In der Interventionsgruppe 1, die sowohl mit internem als auch externem Feedback arbeitete, sowie in der Kontrollgruppe war kein Zuwachs zu verzeichnen (Kapitel 5.2).

Die Ergebnisse stützen die Annahme, dass Feedback positive Effekte auf Lern- und Wissensentwicklung entfalten kann (Hattie & Timperley, 2007; Wisniewski et al., 2019). Allerdings zeigt sich ein differenziertes Bild in Bezug auf die Feedbackquelle. Die besonders hohe Wirksamkeit des internen Feedbacks in Bezug auf das fachdidaktische Wissen lässt sich damit erklären, dass die Selbstbewertung mit dem strukturierten Bewertungsbogen eine hohe metakognitive Aktivierung bei den Proband:innen auslöst. Die intensive Auseinandersetzung mit den Qualitätskriterien und die eigenverantwortliche Beurteilung der selbstentworfenen Lehr-Lern-Materialien erfordert nicht nur eine hohe kognitive Eigenaktivität, sondern auch einen Perspektivenwechsel. Hierbei handelt es sich um Aspekte, die nach Butler und Winne (1995) maßgeblich zur Selbstregulation und zu nachhaltigen Lernprozessen beitragen.

Anders als erwartet, bewirkt das zusätzliche externe Feedback keinen weiteren Wissenszuwachs. Ein potenzieller Grund hierfür ist der erhöhte kognitive Auf-

wand: Die zusätzliche Verarbeitung, der Abgleich und die Reflexion des externen Feedbacks führen zu einem höheren Cognitive Load (Sweller et al., 2011), was dazu führen kann, dass die Potenziale des externen Feedbacks nicht ausgeschöpft werden. Die eigene Selbsteinschätzung steht möglicherweise im Konflikt mit der externen Bewertung, was zu kognitiver Inkonsistenz und Unsicherheit führt (Narciss, 2018).

Das repräsentationsbezogene Fachwissen blieb durch die Intervention vollständig unbeeinflusst. Vermutlich sind die Anforderungen des Bewertungsbogens und der Aufgabenstellung so konzipiert, dass keine inhaltlichen Fachinhalte reflektiert werden müssen. Entsprechende Wissenszuwächse wären eher durch einen fachlich höheren Input zu erwarten. Dieser Befund deckt sich mit den Ergebnissen von Mutke (2017), wonach insbesondere fachdidaktische Wissensbestände sich stärker entwickeln als das eigentliche Fachwissen.

Auffällig ist zudem, dass alle Gruppen bzw. auch die Kontrollgruppe, leichte Fortschritte verzeichnen. Die reine wiederholte Auseinandersetzung mit multiplen externen Repräsentationen und die aktive Gestaltung von Lehr-Lern-Materialien – selbst ohne systematisches Feedback – führen bereits zu positiven Entwicklungen, wie auch bereits in Taskin et al. (2015a) angenommen wurde.

Insgesamt lässt sich festhalten, dass der eigenverantwortliche Einsatz internen Feedbacks eine differenzierte Förderung des fachdidaktischen Wissens, nicht jedoch des Fachwissens anstoßen kann. Sollte externes Feedback zusätzlich angeboten werden, sollte es so angepasst sein, dass es den individuellen Entwicklungsstand sowie die kognitiven Kapazitäten der angehenden Lehrkräfte nicht übersteigt.

6.2.2 F_{1B}: Effekt von internem und/oder externem Feedback auf die Überzeugungen der angehenden Lehrkräfte in Bezug auf die Relevanz von externen Repräsentationen für das Lernen von Chemie

Die Überzeugungen der angehenden Lehrkräfte bezüglich der Relevanz und Nützlichkeit multipler externer Repräsentationen im Chemieunterricht wurden in allen Gruppen als sehr hoch eingeschätzt. Während der gesamten Intervention blieben die Werte weitgehend konstant und es ließen sich keine gruppenspezifischen oder zeitlichen Veränderungen feststellen (Kapitel 5.3).

Die Befunde sind vor dem Hintergrund der bisherigen Forschung plausibel, da Überzeugungen von Lehrkräften im Allgemeinen als zeitlich stabil und veränderungsresistent gelten. Bisherige Interventionsstudien zeigen, dass kurzfristige Maßnahmen, oft nur zu marginalen Veränderungen führen, selbst wenn sie in einem intensiven Maß erfolgen. Nachhaltige Veränderungen hingegen setzen langfristige, praxisnahe und bedeutsame Lernerfahrungen voraus (Fives & Buehl, 2012; Pajares, 1992).

In der hier untersuchten Gruppe liegt vermutlich ein doppelter Verzerrungseffekt vor: Zum einen bringen die angehenden Lehrkräfte aufgrund ihres Masterstudiums inkl. Praxissemesters und Schulpraktika sowie ihrer bisherigen Ausbildung im Vorbereitungsdienst bereits ein ausgeprägtes fachdidaktisches und unterrichtsmethodisches Überzeugungskonstrukt mit. Zum anderen erscheint die sozial erwünschte Bewertung des Testinstruments als nicht ausgeschlossen. Dies ist vor allem möglich, da das Thema der multiplen externen Repräsentationen in der Intervention, wenn auch indirekt, relevant kommuniziert wurde.

Die geringe Veränderungstendenz mag zudem auch darauf zurückzuführen sein, dass die eingesetzten Feedbackformen hauptsächlich kognitive und metakognitive Prozesse und weniger affektive Denkprozesse adressieren (Ryan & Deci, 2000).

6.2.3 F_{1C}: Effekt von internem und/oder externem Feedback auf die Qualität der von den angehenden Lehrkräften designten multiplen externen Repräsentationen

Die Qualität der von den angehenden Lehrkräften selbstdesignten Lehr-Lern-Materialien wurden über den systematisch strukturierten Bewertungsbogen erfasst. Sie lag bereits vor der Intervention und unabhängig von der Interventionsgruppe auf einem bemerkenswert hohen Niveau. Im Verlauf der Intervention gelang es allen Untersuchungsgruppen, das eingesetzte Rohmaterial deutlich zu verbessern. Dabei fiel die Verbesserung punktuell an den Stellen aus, für die im Bewertungsbogen explizite Kriterien benannt waren. Zwischen den Untersuchungsgruppen waren nach der Intervention weder in der Gesamtqualität noch in den einzelnen Kategorien Unterschiede zu beobachten (Kapitel 4.2.5).

Dieses Muster spricht für zwei zentrale Erklärungen: Zum einen verfügen angehende Lehrkräfte im Vorbereitungsdienst auch ohne gezielte Unterstützung bereits über fortgeschrittene Gestaltungskompetenzen. Diese haben sie vermutlich bereits im Studium, in Schulpraktika und bisherigen eigenen Unterrichtsentwürfen entwickelt und routiniert. Zum anderen scheint die wiederholte Bearbeitung von Lehr-Lern-Material hinreichend zu sein, um das bisherige und notwendige Fachwissen und fachdidaktische Wissen zu aktivieren und umzusetzen.

Darüber hinaus weisen die Analysen darauf hin, dass insbesondere die gezielte Ergänzung und Umstrukturierung von Lehr-Lern-Materialien durch strukturiertes internes und externes Feedback schwierig bleibt. Zurückzuführen ist dies auf das bereits fortgeschrittene Repertoire der angehenden Lehrkräfte im Vorbereitungsdienst. Die Intervention scheint daher besonders als Beratung und weniger als Förderung zu wirken. Für weniger erfahrene angehende Lehrkräfte oder Lehramtsstudierende wären daher vermutlich größere Sprünge zu erwarten.

6.2.4 F_{1D}: Effekt von internem und/oder externem Feedback auf die Selbstbewertungsfähigkeit von angehenden Lehrkräften in Bezug auf selbst-designte multiple externe Repräsentationen

Ein zentrales Ergebnis der Interventionsstudie betrifft die Entwicklung der Selbstbewertungsfähigkeit: In beiden Interventionsgruppen 1 und 2 zeigte sich über die drei Sitzungen hinweg eine kontinuierliche Verbesserung der Übereinstimmungswerte zwischen dem internen und externen Feedback. Besonders stark war der Zuwachs in der Interventionsgruppe 1, die sowohl mit internem als auch externem Feedback arbeitete. Die Kontrollgruppe konnte mangels systematischer Daten nicht eingeschätzt werden (Kapitel 4.2.6).

Die Ergebnisse verdeutlichen die Relevanz sowohl systematischen und kriteriengeleiteten als auch externen Feedbacks: Kriteriengeleitetes Feedback ist nicht nur wirksam für das Selbstlernen, sondern auch für die Entfaltung metakognitiver Diagnostik und Reflexionskompetenz (Butler & Winne, 1995; Narciss, 2018). Zudem ist das hier gegebene externe Feedback als externer Real-Wert in der Lage, die Regelgrößen des internen Feedbacks so zu beeinflussen, dass sich der interne Ist-Wert dem Real-Wert angleicht. Dies führt zu einer sukzessiven Verminderung der Diskrepanzen zwischen den internen und externen Bewertungen und einem Aufbau der Selbstbewertungsfähigkeit.

Bemerkenswert ist zudem die Dynamik der Bewertungstendenz: In den ersten Sitzungen schätzten die angehenden Lehrkräfte ihre designten Lehr-Lern-Materialien über oder unter der tatsächlichen Bewertung (externes Feedback) ein. Mit zunehmender Wiederholung der Feedbackschleife näherte sich die Einschätzung der angehenden Lehrkräfte an das externe Rating an. Dies gilt insbesondere für die Qualität der Abbildung, des Texts und der Repräsentationsebenen. Für fachdidaktisch bekannte Bereiche war die Selbsteinschätzung tendenziell zu niedrig (z. B. Lernziel, inhaltliche Passung, Einleitung in die Thematik), für neue Inhalte hingegen eher zu hoch (z. B. Abbildung, Repräsentationsebenen). Dies bildet ein klassisches Muster im Entwicklungsverlauf metakognitiver Expertise ab (vgl. Hattie & Timperley, 2007).

Die Wahl einer differenzierten fünfstufigen Likert-Skala erwies sich als valide und als lernförderlich. Dies gilt vor allem für die Interventionsgruppe 2, die sowohl mit internem als auch externem Feedback arbeitete. Durch die Verkleinerung der Skalenbreite hat sich die Selbsteinschätzungsfähigkeit nicht verbessert. Umgekehrte Effekte waren jedoch bei der Interventionsgruppe 1 zu finden, die mit einer dreistufigen Skala besser zurechtgekommen wären. Zusammengefasst lässt sich daher auch sagen, dass das zusätzliche externe Feedback die angehenden Lehrkräfte zu einer differenzierteren und präziseren Selbsteinschätzung befähigt.

In Bezug auf die unterrichtspraktische Professionalisierung angehender Lehrkräfte befürworten die Erkenntnisse daher die Kombination und den Abgleich von Selbstfeedbacks mit tutoriellem Fremdfeedback.

6.2.5 *F₂: Effekt der Stichprobenmerkmale der angehenden Lehrkräfte auf das repräsentationsbezogene Fachwissen und fachdidaktische Wissen*

Die explorativen Analysen zur Bedeutung personenbezogener Voraussetzungen bestätigten, dass die Lehramtsform und Zugangsqualifikation der angehenden Lehrkräfte das Fachwissen signifikant beeinflussen. Die angehenden Gymnasiallehrkräfte erzielten durchweg höhere Leistungen sowohl im allgemeinen als auch repräsentationsbezogenen Fachwissen. Im fachdidaktischen Wissen zeigten sich die Differenzen weniger stark. Es ergaben sich zudem signifikante Korrelationen zwischen allgemeinem und repräsentationsbezogenem Fachwissen, während Korrelationen zwischen fachdidaktischem Wissen und fachlichem

Vorwissen schwächer ausgeprägt waren. Darüber hinaus war das repräsentationsbezogene Vorwissen ein zentraler Prädiktor für den repräsentationsbezogenen Lernzuwachs. Andere Personenmerkmale wie Abschlussnoten, Zugangsqualifikation oder persönliche Ressourcen spielten dagegen eine geringere Rolle (Kapitel 5.6).

Diese Resultate bestätigen die Ergebnisse von Dollny (2011), Mutke (2017) und Taskin et al. (2015a), nach denen die Ausbildung an unterschiedlichen Schulformen in Chemie deutliche Auswirkungen auf das Wissen der angehenden Lehrkräfte haben. Es kann festgehalten werden, dass die feedbackbasierte Intervention Teile der professionellen Kompetenz angehender Lehrkräfte fördern, aber bestehende Leistungsunterschiede nicht ausgleichen kann. Hier bedarf es zusätzlicher, gestufter und passgenauer Unterstützungsmaßnahmen.

6.3 Gesamtfazit, Implikationen für die Forschung und Schulpraxis

Die vorliegende Studie untersuchte die differenziellen Effekte internen und externen Feedbacks auf die professionelle Kompetenz angehender Chemielehrkräfte im Vorbereitungsdienst unter besonderer Berücksichtigung der Gestaltung und Bewertung multipler externer Repräsentationen. Durch das entwickelte Studiendesign, die vielfältigen Testinstrumente und die Einbindung der Intervention in den Praxiskontext wurden zentrale Aspekte der aktuellen Professionalisierungsforschung für Lehrkräfte systematisch adressiert.

Im Gesamtbild lassen sich mehrere übergeordnete Erkenntnisse festhalten: Erstens zeigt sich, dass feedbackgestützte Interventionen positive, aber selektive Effekte auf verschiedene Bereiche der professionellen Kompetenz ausüben. So konnte durch den wiederholten Einsatz des kriteriengeleiteten internen Feedbacks insbesondere das repräsentationsbezogene fachdidaktische Wissen der angehenden Lehrkräfte signifikant gesteigert werden. Die Kombination aus internem und externem Feedback führte dagegen zu einer signifikanten Verbesserung der Selbstbewertungsfähigkeit.

Zweitens verdeutlichen die Ergebnisse, dass die repräsentationsbezogenen Überzeugungen angehender Lehrkräfte bereits auf einem hohen Niveau ausgebildet sind und im untersuchten Interventionszeitraum kaum veränderbar erscheinen. Hier spielt vermutlich die zeitliche Stabilität solcher Überzeugungen sowie ein möglicher Deckeneffekt der eingesetzten Messinstrumente eine Rolle. Zukünftig wäre daher zu prüfen, ob längerfristig angelegte Maßnahmen stärkere Veränderungen ermöglichen.

Drittens wurde deutlich, dass die Qualität der selbst designten Lehr-Lern-Materialien nicht durch mangelnde Expertise der angehenden Lehrkräfte gehemmt wird. Ein bereits hohes Ausgangsniveau der angehenden Lehrkräfte limitierte lediglich ihr Verbesserungspotenzial. Demnach wären stärkere Effekte vor allem bei weniger erfahrenen angehenden Lehrkräften oder Lehramtsstudierenden zu erwarten.

Praktisch bedeuten diese Ergebnisse, dass die Lehrkräftebildung im Vorbereitungsdienst vom Einsatz kriteriengeleiteter Feedbackinstrumente profitiert, insbesondere wenn diese sowohl zur Selbsteinschätzung als auch zur Fremdbeurteilung genutzt werden. Die Stärkung der Selbstbewertungsfähigkeit

ist hier als grundlegende Basis für kontinuierliche professionelle Weiterentwicklung anzusehen. Zudem legen die Ergebnisse nahe, dass Feedback nicht als einmaliges Ereignis, sondern als fester Bestandteil in zyklische Entwicklungsprozesse mit expliziten Reflexionsphasen integriert werden sollte. Für die Praxis bedeutet dies unter anderem:

- die Verankerung kriterienbasierter Bewertungsbögen und Feedbackprozesse in die regulären Ausbildungsabschnitte des Vorbereitungsdienstes zu fördern,
- angehende Lehrkräfte gezielt zum metakognitiven Abgleich von Eigen- und Fremdeinschätzung anzuleiten,
- zusätzliche Unterstützungsangebote für weniger erfahrene oder unsichere angehende Lehrkräfte bereitzustellen, um binnendifferenzierte Lernerfolge zu ermöglichen.

Insgesamt leisten die vorliegenden Ergebnisse einen fundierten Beitrag zur evidenzbasierten Weiterentwicklung der Lehrkräftebildung in der zweiten Phase und bieten praxisnahe Anregungen zur Stärkung der professionellen Kompetenz zukünftiger Chemielehrkräfte.

6.4 Limitationen und Ausblick

Im Rahmen dieser empirischen Arbeit konnten zentrale Effekte internen und externen Feedbacks auf die professionelle Kompetenz angehender Chemielehrkräfte aufgezeigt werden. Dennoch sind einige Limitationen in methodischer, inhaltlicher und kontextueller Hinsicht zu berücksichtigen, die für die Interpretation und Generalisierung der Befunde bedeutsam sind.

Erstens ist die Zusammensetzung und Größe der Stichprobe zu beachten. Obwohl eine vergleichsweise große und heterogene Gruppe angehender Lehrkräfte im Vorbereitungsdienst erreicht werden konnte, handelt es sich dennoch um eine spezifische Auswahl aus dem Bundesland Nordrhein-Westfalen mit fachspezifischem Fokus auf Chemielehrkräfte. Die Übertragbarkeit der Ergebnisse auf andere Bundesländer, Fächer sowie auf allgemeinbildende oder berufsbildende Schulen bleibt somit eingeschränkt. Auch war die Dropout-Quote im Verlauf der Online-Intervention trotz motivierender Maßnahmen nicht unerheblich, was sowohl auf externe Einflussfaktoren (z. B. pandemiebedingte Belastungen) als auch auf die Eigenverantwortung der Teilnehmenden im asynchronen Setting zurückzuführen ist (vgl. Döring & Bortz, 2023).

Zweitens zeigen sich im Bereich der eingesetzten Testinstrumente trotz intensiver Pilotierung und Güteprüfung gewisse Einschränkungen hinsichtlich der Reliabilität, insbesondere bei den repräsentationsbezogenen Skalen. Zudem können sozial erwünschte Antworttendenzen insbesondere in Selbsteinschätzungen nie ganz ausgeschlossen werden.

Drittens liegt eine methodische Einschränkung im Interventionsdesign selbst: Die Umsetzung in einem digitalisierten, asynchronen Online-Format war einerseits notwendig, brachte aber andererseits Nachteile hinsichtlich der direkten Steuerbarkeit der Lernsituation und der Überprüfbarkeit der tatsächlichen Bearbeitungstiefe mit sich.

Viertens ist das Studiendesign bewusst auf die Differenzierung von Feedbackformen und die Entwicklung professioneller Kompetenz fokussiert, während andere relevante Dimensionen wie der Transfer auf die tatsächliche Unterrichtspraxis oder Schüler:innenperspektiven nicht erfasst wurden. Ebenso wurde der Beobachtungszeitraum auf wenige Wochen begrenzt, sodass sich keine Aussagen zum Langzeiteffekt oder zur nachhaltigen Integration in das professionelle Handeln treffen lassen.

Ausblickend ergeben sich aus den genannten Limitationen zahlreiche weiterführende Forschungsfelder. Zukünftige Arbeiten sollten beispielsweise gezielte Varianten von Peer-Feedback oder kollegialem Coaching analysieren, um zusätzliche Unterstützungspotenziale zu erschließen (Wisniewski et al., 2019). Eine längsschnittliche Untersuchung über mehrere Phasen der Lehrkräftebildung hinweg wäre ebenso nützlich, um die Nachhaltigkeit von Feedbackinterventionen und Überzeugungsveränderungen genauer untersuchen zu können.

Weiterhin empfiehlt sich die Entwicklung und Validierung sensiblerer Messinstrumente zur Erfassung von fachbezogenem Professionswissen und Überzeugungen. Die Erweiterung der vorliegenden Studie auf andere naturwissenschaftliche oder auch gesellschaftswissenschaftliche Unterrichtsfächer kann zudem helfen, fachspezifische und domänenübergreifende Lernmechanismen zu differenzieren (Kleickmann et al., 2014).

Abschließend ist festzuhalten, dass die vorliegende Studie trotz der genannten Einschränkungen einen wesentlichen Beitrag zum Verständnis von Feedbackprozessen in der zweiten Phase der Lehrkräfteausbildung leistet. Die identifizierten Limitationen dienen zugleich als Wegweiser für nachfolgende Arbeiten und als Anstoß für eine noch gezieltere, individuelle und differenzierte Professionalisierung künftiger Lehrkräfte.

7 VERZEICHNISSE

7.1 Literaturverzeichnis

- Abs, H. J., Döbrich, P., Vögele, E. & Klieme, E. (2005). *Skalen zur Qualität der Lehrerbildung - Dokumentation der Erhebungsinstrumente: Pädagogische Entwicklungsbilanzen an Studienseminaren (PEB-Sem)* (2. Aufl.). *Materialien zur Bildungsforschung: Bd. 12*.
- Adamina, M. (2014). Lehr- und Lernmaterialien im kompetenzorientierten Unterricht. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 32(3), 359–372.
<https://doi.org/10.25656/01:13875>
- Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, 33(2-3), 131–152.
[https://doi.org/10.1016/S0360-1315\(99\)00029-9](https://doi.org/10.1016/S0360-1315(99)00029-9)
- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183–198.
<https://doi.org/10.1016/j.learninstruc.2006.03.001>
- Ainsworth, S. (2018). Multiple Representations and Multimedia Learning. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman & P. Reimann (Hrsg.), *International handbook of the learning sciences* (S. 96–105). Routledge.
- Anderson-Park, E. & Abs, H. J. (2020). Lehrerinnen- und Lehrerbildung im Vorbereitungsdienst. In C. Cramer, J. König, M. Rothland & S. Blömeke (Hrsg.), *Handbuch Lehrerinnen- und Lehrerbildung*. Verlag Julius Klinkhardt.
<https://doi.org/10.35468/hblb2020-038>
- Asselborn, W., Jäckel, M. & Risch, K. T. (Hrsg.). (2004). *Chemie heute - Sekundarbereich I: Gesamtband - Nordrhein-Westfalen*. Schroedel.
- Asselborn, W., Jäckel, M. & Risch, K. T. (Hrsg.). (2014). *Chemie heute: Sekundarstufe I. Nordrhein-Westfalen* (1. Aufl.). Schroedel.
- Baddeley, A. D., Eysenck, M. W. & Anderson, M. C. (2015). *Memory* (2. ed.). Psychology Press.
- Ballstaedt, S.-P. (1997). *Wissensvermittlung: Die Gestaltung von Lernmaterial*. Beltz Psychologie Verlags Union.

- Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A. & Morgan, M. (1991). The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research*, 61(2), 213–238.
<https://doi.org/10.3102/00346543061002213>
- Barke, H.-D. & Harsch, G. (2011). *Chemiedidaktik kompakt: Lernprozesse in Theorie und Praxis*. Springer Berlin Heidelberg.
<https://doi.org/10.1007/978-3-64220220-9>
- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Waxmann.
- Behnke, K. (2016). Umgang mit Feedback bei zwei ausgewählten Personen-
gruppen des schulischen Miteinanders. In K. Behnke (Hrsg.), *Umgang mit Feedback im Kontext Schule* (S. 43–66). Springer Fachmedien Wiesbaden.
https://doi.org/10.1007/978-3-658-10223-4_4
- Behrens, D., Forell, M., Idel, T.-S. & Pauling, S. (Hrsg.). (2023). *Studien zur Professionsforschung und Lehrer:innenbildung. Lehrkräftebildung in der Bedarfskrise: Programme – Positionierungen – Empirie*. Verlag Julius Klinkhardt.
- Bennett, J., Lubben, F. & Hogarth, S. (2006). Bringing science to life: A synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. *Science Education*, 91(3), 347–370.
<https://doi.org/10.1002/sce.20186>
- Bitan, K. (2014). *Umgang mit Feedback am Beispiel Referendariat und Schulentwicklung – eine sozialpsychologische Analyse: Dissertation*.
https://duepublico2.uni-due.de/receive/duepublico_mods_00036425
- Blanz, M. (2021). *Forschungsmethoden und Statistik für die Soziale Arbeit: Grundlagen und Anwendungen* (2. Auflage). W. Kohlhammer Verlag.
<https://doi.org/10.17433/978-3-17-039819-1>
- Bleicher, R. E. (2004). Revisiting the STEBI-B: Measuring Self-Efficacy in Pre-service Elementary Teachers. *School Science and Mathematics*, 104(8), 383–391.
<https://doi.org/10.1111/j.1949-8594.2004.tb18004.x>

- Bodner, G. M. & Domin, D. S. (2000). Mental Models: The Role of Representations in Problem Solving in Chemistry. *UNIVERSITY CHEMISTRY EDUCATION*, 4(1), 24–30.
- Brinko, K. T. (1993). The Practice of Giving Feedback to Improve Teaching: What Is Effective? *The Journal of Higher Education*, 64(5), 574.
<https://doi.org/10.2307/2959994>
- Brown, T. L., LeMay, H. E., Bursten, B. E. & Robl, C. (2007). *Chemie: Die zentrale Wissenschaft* (Chemie - Studieren kompakt). *che, Chemie*. Pearson Studium.
- Bühner, M. (2021). *Einführung in die Test- und Fragebogenkonstruktion* (4., korrigierte und erweiterte Auflage). *Pearson Studium Psychologie*. Pearson.
<https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6741201>
- Butler, D. L. & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, 65(3), 245–281.
<https://doi.org/10.3102/00346543065003245>
- Cheng, M. & Gilbert, J. K. (2009). Towards a Better Utilization of Diagrams in Research into the Use of Representative Levels in Chemical Education. In J. K. Gilbert & D. F. Treagust (Hrsg.), *Multiple Representations in Chemical Education* (S. 55–73). Springer Netherlands.
- Clark, J. M. & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149–210.
<https://doi.org/10.1007/BF01320076>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Taylor and Francis.
<https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=1192162>
- Cromley, J. G., Snyder-Hogan, L. E. & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology*, 35(1), 59–74.
<https://doi.org/10.1016/j.cedpsych.2009.10.002>

- Derman, A. & Ebenezer, J. (2020). The Effect of Multiple Representations of Physical and Chemical Changes on the Development of Primary Pre-service Teachers Cognitive Structures. *Research in Science Education*, 50(4), 1575–1601.
<https://doi.org/10.1007/s11165-018-9744-5>
- diSessa, A. A. (2004). Metarepresentation: Native Competence and Targets for Instruction. *Cognition and Instruction*, 22(3), 293–331.
https://doi.org/10.1207/s1532690xci2203_2
- Ditton, H. & Müller, A. (2014). Einleitung. In H. Ditton (Hrsg.), *Feedback und Rückmeldungen: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 7–10). Waxmann.
- Dollny, S. (2011). *Entwicklung und Evaluation Eines Testinstruments Zur Erfassung des Fachspezifischen Professionswissens Von Chemielehrkräften. Studien Zum Physik- und Chemielernen Ser: v.127*. Logos Verlag Berlin.
<https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5219703>
- Döring, N. & Bortz, J. (2023). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (6. Aufl.). Springer.
- Draxler, D. (2005). *Aufgabendesign und basismodellorientierter Physikunterricht: Dissertation*.
<https://nbn-resolving.org/urn:nbn:de:hbz:465-20070412-082954-2>
- Dubberke, T., Kunter, M., McElvany, N., Brunner, M. & Baumert, J. (2008). Lerntheoretische Überzeugungen von Mathematiklehrkräften. *Zeitschrift für Pädagogische Psychologie*, 22(34), 193–206.
<https://doi.org/10.1024/1010-0652.22.34.193>
- Eilam, B. & Poyas, Y. (2008). Learning with multiple representations: Extending multimedia learning beyond the lab. *Learning and Instruction*, 18(4), 368–378.
<https://doi.org/10.1016/j.learninstruc.2007.07.003>
- Eisner, W., Gietz, P., Justus, A., Laitenberger, K., Nickolay, H., Schierle, W., Schmidt, B., Sternberg, M. & Zippel, T. (2009). *Elemente Chemie 1: Klasse 7-9*. Nordrhein-Westfalen - G8 (1. Aufl.).
- Ernst, C. & Becker, F.-M. (Hrsg.). (2009). *Chemie: Gesamtband Sekundarstufe I* (1. Aufl.). Duden-Paetec-Schulbuchverl.

- Field, A. (2024). *Discovering statistics using IBM SPSS statistics* (6th edition). Sage.
- Fives, H. & Buehl, M. M. (2012). Spring cleaning for the “messy” construct of teachers’ beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer & M. Zeidner (Hrsg.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors* (S. 471–499). American Psychological Association.
<https://doi.org/10.1037/13274-019>
- Fluck, H.-R. (1996). *Fachsprachen: Einführung und Bibliographie* (5., überarb. und erw. Aufl.). *UTB Germanistik, Linguistik: Bd. 483*. Francke.
- Gehrmann, A. (2023). Quer- und Seiteneinstiege in den Lehrer:innenberuf – Gründe, Spielarten und Folgen alternativer Wege in die Schule. In D. Behrens, M. Forell, T.-S. Idel & S. Pauling (Hrsg.), *Studien zur Professionsforschung und Lehrer:innenbildung. Lehrkräftebildung in der Bedarfskrise: Programme - Positionierungen - Empirie* (S. 25–53). Verlag Julius Klinkhardt.
- Gietz, P., Jung, U., Knetsch, R., Lehmacher, W., Peppmeier, R., Schäpers, B., Wald-Schillings, I., Weizel, B. & Zetzel, S. (2013). *Prisma Chemie: Differenzierte Ausgabe mit CD-ROM* (1. Aufl.). *Ausgabe A*. Klett.
- Gilbert, J. K. & Treagust, D. F. (2009a). Macro, Submicro and Symbolic Representations and the Relationship Between Them: Key Models in Chemical Education. In J. K. Gilbert & D. F. Treagust (Hrsg.), *Multiple Representations in Chemical Education* (S. 1–8). Springer Netherlands.
- Gilbert, J. K. & Treagust, D. F. (Hrsg.). (2009b). *Multiple Representations in Chemical Education* (Bd. 4). Springer Netherlands.
<https://doi.org/10.1007/978-1-4020-8872-8>
- Gilbert, J. K. & Treagust, D. F. (2009c). Towards a Coherent Model for Macro, Submicro and Symbolic Representations in Chemical Education. In J. K. Gilbert & D. F. Treagust (Hrsg.), *Multiple Representations in Chemical Education* (S. 333–350). Springer Netherlands.
- Grosjean, F. (2021). *Life As a Bilingual: Knowing and Using Two or More Languages* (1st ed.). Cambridge University Press.
<https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6868806>

- Güldener, T. (2023). MINT-Lehrkräfte gesucht – wo bleibt der Nachwuchs? In D. Behrens, M. Forell, T.-S. Idel & S. Pauling (Hrsg.), *Studien zur Professionsforschung und Lehrer:innenbildung. Lehrkräftebildung in der Bedarfskrise: Programme - Positionierungen - Empirie* (S. 172–190). Verlag Julius Klinkhardt.
- 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.
<https://doi.org/10.1037//0022-0663.90.3.414>
- Hascher, T. (2014). Forschung zur Wirksamkeit der Lehrerbildung. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 418–440). Waxmann.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112.
<https://doi.org/10.3102/003465430298487>
- Hattie, J. & Wollenschläger, M. (2014). A conceptualization of feedback. In H. Ditton (Hrsg.), *Feedback und Rückmeldungen: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 135–150). Waxmann.
- Hattie, J. & Zierer, K. (2025). *10 Mindframes for Visible Learning* (2. Aufl.). Routledge.
<https://doi.org/10.4324/9781003430124>
- Hoffmann, R. & Laszlo, P. (1991). Representation in Chemistry. *Angewandte Chemie International Edition in English*, 30(1), 1–16.
<https://doi.org/10.1002/anie.199100013>
- IBM SPSS Statistics for Windows (Version 23.0) [Computer software]. (2016). IBM Corp.
<https://www.ibm.com/us-en/?lnk=m>
- Kagan, D. M. (1992). Professional Growth among Preservice and Beginning Teachers. *Review of Educational Research*, 62(2), 129.
<https://doi.org/10.2307/1170578>
- Kärner, T., Bonnes, C. & Schölzel, C. (2019). Bewertungstransparenz im Referendariat. Vorab-Onlinepublikation.
<https://doi.org/10.25656/01:23948>

- Keller-Schneider, M. & Hericks, U. (2011). Forschung zum Berufseinstieg: Übergang von der Ausbildung in den Beruf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 297–313). Waxmann.
- Kleickmann, T. & Anders, Y. (2011). Lernen an der Universität. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 305–315). Waxmann.
- Kleickmann, T., Großschedl, J., Harms, U., Heinze, A., Herzog, S., Hohenstein, F., Köller, O., Kröger, J., Lindmeier, A., Loch, C., Mahler, D., Möller, J., Neumann, K., Parchmann, I., Steffensky, M., Taskin, V. & Zimmermann, F. (2014). Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer – Testentwicklung im Rahmen des Projekts KiL. *Unterrichtswissenschaft*, 42(3), 280–288.
- Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S. & Baumert, J. (2013). Teachers' Content Knowledge and Pedagogical Content Knowledge. *Journal of Teacher Education*, 64(1), 90–106.
<https://doi.org/10.1177/0022487112460398>
- Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284.
<https://doi.org/10.1037/0033-2909.119.2.254>
- Klusmann, U. (2011). Individuelle Voraussetzungen von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 295–304). Waxmann.
- Koenen, J. (2014). *Entwicklung und Evaluation von experimentunterstützten Lösungsbeispielen zur Förderung naturwissenschaftlich-experimenteller Arbeitsweisen. Studien zum Physik- und Chemielernen: Bd. 171*. Logos.
- Kozma, R. & Russell, J. (1997). Multimedia and Understanding: Expert and Novice Responses to Different Representations of Chemical Phenomena. *Journal of Research in Science Teaching*, 34(9), 949–968.
[https://doi.org/10.1002/\(SICI\)1098-2736\(199711\)34:9<949::AID-TEA7>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1098-2736(199711)34:9<949::AID-TEA7>3.0.CO;2-U)

- Kozma, R. & Russell, J. (2005). Students Becoming Chemists: Developing Representationl Competence. In J. K. Gilbert (Hrsg.), *Models and Modeling in Science Education: Bd. 1. Visualization in science education* (S. 121–145). Springer.
https://doi.org/10.1007/1-4020-3613-2_8
- Krey, O. & Schwanewedel, J. (2018). Lernen mit externen Repräsentationen. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Theorien in der naturwissenschaftsdidaktischen Forschung* (Bd. 33, S. 159–175). Springer Berlin Heidelberg.
https://doi.org/10.1007/978-3-662-56320-5_10
- Kulhavy, R. W. & Stock, W. A. (1996). How Cognitive Maps are Learned and Remembered. *Annals of the Association of American Geographers*, 86(1), 123–145.
<https://doi.org/10.1111/j.1467-8306.1996.tb01748.x>
- Rahmenvereinbarung über die Ausbildung und Prüfung für ein Lehramt der Sekundarstufe I (Lehramtstyp 3) (Beschluss der Kultusministerkonferenz vom 28.02.1997 i. d. F. vom 13.09.2018) (1997 & i.d.F.v. 2018).
https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluess e/1997/1997_02_28-RV-Lehramtstyp-3.pdf
- Rahmenvereinbarung über die Ausbildung und Prüfung für ein Lehramt der Sekundarstufe II (allgemeinbildende Fächer) oder für das Gymnasium (Lehramtstyp 4) (Beschluss der Kultusministerkonferenz vom 28.02.1997 i. d. F. vom 13.09.2018) (1997 & i.d.F.v. 2018).
https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1997/1997_02_28-RV_Lehramtstyp-4_.pdf
- Weiterentwickelte Bildungsstandards in den Naturwissenschaften für das Fach Chemie (MSA) (Beschluss der Kultusministerkonferenz vom 16.12.2004 i.d.F. vom 13.06.2024) (2004).
https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluess e/2024/2024_06_13-WeBiS_Chemie_MSA.pdf
- Standards für die Lehrerbildung: Bildungswissenschaften (2004).
https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf

- Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung (Beschluss der Kultusministerkonferenz vom 16.10.2008 i. d. F. vom 08.02.2024) (2008 & i.d.F.v. 2024).
https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluess/2008/2008_10_16-Fachprofile-Lehrerbildung.pdf
- Ländergemeinsame Anforderungen für die Ausgestaltung des Vorbereitungsdienstes und die abschließende Staatsprüfung (Beschluss der Kultusministerkonferenz vom 06.12.2012) (2012).
https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2012/2012_12_06-Vorbereitungsdienst.pdf
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Waxmann.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T. & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820.
<https://doi.org/10.1037/a0032583>
- Lachmayer, S., Nerdel, C. & Prechtel, H. (2007). Modellierung kognitiver Fähigkeiten beim Umgang mit Diagrammen im naturwissenschaftlichen Unterricht: Modelling of cognitive abilities regarding the handling of graphs in science education. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 145–160.
- Leisen, J. (1998). Förderung des Sprachlernens durch den Wechsel von Symbolisierungsformen im Physikunterricht. *Praxis der Naturwissenschaften Physik*, 47(2), 9–13.
- Leisen, J. (2005). Wechsel der Darstellungsformen: Eine wichtige Strategie im kommunikativen Physikunterricht. *Unterricht Physik*, 16(87), 10–11.
- Lenhard, W. & Lenhard, A. (2017). *Computation of Effect Sizes*.
<https://doi.org/10.13140/RG.2.2.17823.92329>
- Levie, W. H. & Lentz, R. (1982). Effects of text illustrations: A review of research. *ECTJ*, 30(4), 195–232.
<https://doi.org/10.1007/BF02765184>

- Lipnevich, A. A. & Panadero, E. (2021). A Review of Feedback Models and Theories: Descriptions, Definitions, and Conclusions. *Frontiers in Education*, 6, Artikel 720195.
<https://doi.org/10.3389/feduc.2021.720195>
- Lipowsky, F. (2009). Unterrichtsentwicklung durch Fort- und Weiterbildungsmaßnahmen für Lehrpersonen. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 27(3), 346–360.
- Lipowsky, F. (2011). Theoretische Perspektiven und empirische Befunde zur Wirksamkeit von Lehrerfort- und -weiterbildung. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 398–417). Waxmann.
- Magnus, L., Schütte, K. & Schwanewedel, J. (2020). Challenges Solving Science Tasks with Text–Picture Combinations Persist beyond Secondary School. *Journal of Research on Educational Effectiveness*, 13(4), 759–783.
<https://doi.org/10.1080/19345747.2020.1750744>
- Mangold, R. (2015). *Informationspsychologie: Wahrnehmen und Gestalten in der Medienwelt* (2. Aufl.). *Lehrbuch*. Springer.
- Mayer, R. E. (Hrsg.). (2009a). *Multimedia learning* (2. Aufl.). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511811678>
- Mayer, R. E. (2009b). The Science of Instruction: Determining What Works in Multimedia Learning. In R. E. Mayer (Hrsg.), *Multimedia learning* (2. Aufl., S. 28–56). Cambridge University Press.
- Mayer, R. E. (2009c). The Science of Instruction: Determining What Works in Multimedia Learning. In R. E. Mayer (Hrsg.), *Multimedia learning* (2. Aufl., S. 28–56). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511811678.004>
- Mayer, R. E. (2009d). The Science of Learning: Determining How Multimedia Learning Works. In R. E. Mayer (Hrsg.), *Multimedia learning* (2. Aufl., S. 57–83). Cambridge University Press.
- Mayer, R. E. (Hrsg.). (2014a). *The Cambridge handbook of multimedia learning* (2. Aufl.). Cambridge University Press.

- Mayer, R. E. (2014b). Introduction to Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (2. Aufl., S. 1–24). Cambridge University Press.
- McElvany, N., Schroeder, S., Baumert, J., Schnotz, W., Horz, H. & Ullrich, M. (2012). Cognitively demanding learning materials with texts and instructional pictures: teachers' diagnostic skills, pedagogical beliefs and motivation. *European Journal of Psychology of Education*, 27(3), 403–420.
<https://doi.org/10.1007/s10212-011-0078-1>
- McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W., Horz, H. & Ullrich, M. (2009). Diagnostische Fähigkeiten von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, 23(34), 223–235.
<https://doi.org/10.1024/1010-0652.23.34.223>
- McElvany, N., Schroeder, S., Richter, T., Hachfeld, A., Baumert, J., Schnotz, W., Horz, H. & Ullrich, M. (2010). Texte mit instruktionalen Bildern als Unterrichtsmaterial – Kompetenzen der Lehrkräfte. *Unterrichtswissenschaft*, 38(2), 98–116.
- Mertner, N. O. (2021). *Entwicklung und Evaluation einer Unterstützungsmaßnahme zur Erstellung von multiplen externen Repräsentationen für angehende Lehrkräfte*. Masterarbeit. Universität Duisburg-Essen, Didaktik der Chemie.
- Milenković, D. D., Segedinac, M. D. & Hrin, T. N. (2014). Increasing High School Students' Chemistry Performance and Reducing Cognitive Load through an Instructional Strategy Based on the Interaction of Multiple Levels of Knowledge Representation. *Journal of Chemical Education*, 91(9), 1409–1416.
<https://doi.org/10.1021/ed400805p>
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. (2015). *Chemie - Übersicht über die Operatoren*.
<https://www.standardsicherung.schulministerium.nrw.de/cms/zentralabitur-gost/faecher/getfile.php?file=3850>

- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (Hrsg.). (2021). *Kerncurriculum für die Lehrerbildung im Vorbereitungsdienst: Verbindliche Zielvorgabe der schulpraktischen Lehrerbildung in Nordrhein-Westfalen*.
https://www.schulministerium.nrw/system/files/media/document/file/Kerncurriculum_Vorbereitungsdienst.pdf
- Gesetz über die Ausbildung für Lehrämter an öffentlichen Schulen (Lehrerbildungsgesetz - LABG) (2022).
https://recht.nrw.de/lmi/owa/br_bes_text?anw_nr=2&bes_id=12764
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (Hrsg.). (2008). *Schule in NRW: Bd. 3415. Kernlehrplan für das Gymnasium - Sekundarstufe I in Nordrhein-Westfalen: Chemie* (1. Aufl.). Ritterbach.
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (Hrsg.). (2011a). *Schule in NRW: Bd. 3204. Kernlehrplan für die Hauptschule in Nordrhein-Westfalen: Lernbereich Naturwissenschaften. Biologie, Chemie, Physik* (1. Aufl.). Ritterbach.
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (Hrsg.). (2011b). *Schule in NRW: Bd. 3308. Kernlehrplan für die Realschule in Nordrhein-Westfalen: Chemie* (1. Aufl.). Ritterbach.
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (Hrsg.). (2013). *Schule in NRW: Bd. 3108. Kernlehrplan für die Gesamtschule - Sekundarstufe I in Nordrhein-Westfalen: Naturwissenschaften, Biologie, Chemie, Physik* (2. Aufl.). Ritterbach.
- Mühlhaus, L. (2020). *Evaluation eines moodle-basierten Moduls zur Bewertung von Lehr-Lern-Materialien für den Chemieunterricht: Masterarbeit*. Universität Duisburg-Essen, Didaktik der Chemie.
- Müller, A. & Ditton, H. (2014). Feedback: Begriff, Formen und Funktionen. In H. Ditton (Hrsg.), *Feedback und Rückmeldungen: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 11–28). Waxmann.
- Müller, W., Fischer, H. E. & Duit, R. (2002). Vielfalt und Anregung statt Routine. *Unterricht Physik*(67), 4–7 (Aufgaben).

- Mutke, S. (2017). *Das Professionswissen von Chemiereferendarinnen und -referendaren in Nordrhein-Westfalen: Eine Längsschnittstudie*. Dissertation. *Studien zum Physik- und Chemielernen: Bd. 222*. Logos.
- Narciss, S. (2006). *Informatives tutorielles Feedback: Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse*. *Pädagogische Psychologie und Entwicklungspsychologie: Bd. 56*. Waxmann.
- Narciss, S. (2014). Modelle zu den Bedingungen und Wirkungen von Feedback in Lehr-Lernsituationen. In H. Ditton (Hrsg.), *Feedback und Rückmeldungen: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 43–82). Waxmann.
- Narciss, S. (2018). Feedbackstrategien für interaktive Lernaufgaben. In S. Kracht, A. Niedostadek & P. Sensburg (Hrsg.), *Springer eBook Collection. Praxishandbuch Professionelle Mediation: Methoden, Tools, Marketing und Arbeitsfelder* (Living reference work, S. 1–24). Springer.
https://doi.org/10.1007/978-3-662-54373-3_35-1
- Naroll, R. (1974). An Exact Test of Significance for Goodman and Kruskal's Gamma. *Behavior Science Research*, 9(1), 27–40.
<https://doi.org/10.1177/106939717400900112>
- Nerdel, C. (2017). *Grundlagen der Naturwissenschaftsdidaktik: Kompetenzorientiert und aufgabenbasiert für Schule und Hochschule*. Springer.
<https://doi.org/10.1007/978-3-662-53158-7>
- Nitz, S. (2012). *Fachsprache im Biologieunterricht: Eine Untersuchung zu Bedingungsfaktoren und Auswirkungen* [Dissertation]. Christian-Albrechts-Universität zu Kiel, Kiel.
https://macau.uni-kiel.de/servlets/MCRFileNodeServlet/dissertation_derivate_00004276/diss_nitz.pdf
- Nitz, S., Enzingmüller, C., Precht, H. & Nerdel, C. (2011). Fachsprache im naturwissenschaftlichen Unterricht - eine empirische Untersuchung zur Einstellung angehender Lehrkräfte. *Unterrichtswissenschaft*, 39(3), 245–262.

- Opfermann, M., Schmeck, A. & Fischer, H. E. (2017). Multiple Representations in Physics and Science Education – Why Should We Use Them? In D. F. Treagust, R. Duit & H. E. Fischer (Hrsg.), *Multiple Representations in Physics Education* (S. 1–22). Springer.
- Oser, F. & Blömeke, S. (2012). Überzeugungen von Lehrpersonen: Einführung in den Thementeil. *Zeitschrift für Pädagogik*, 58(4), 415–421.
- Pajares, M. F. (1992). Teachers' Beliefs and Educational Research: Cleaning Up a Messy Construct. *Review of Educational Research*, 62(3), 307–332.
<https://doi.org/10.3102/00346543062003307>
- Permatasari, M. B., Rahayu, S. & Dasna, I. W. (2022). Chemistry Learning Using Multiple Representations: A Systematic Literature Review. *Journal of Science Learning*, 5(2), 334–341.
<https://doi.org/10.17509/jsl.v5i2.42656>
- Peterson, D. (1996). *Forms of representation: An interdisciplinary theme for cognitive science*. Intellect.
- Peukert, L. (2018). *Entwicklung und Validierung eines Messinstrumentes zur Erhebung der Überzeugungen von angehenden Lehrkräften zu externen Repräsentationen*. Masterarbeit. Universität Duisburg-Essen, Didaktik der Chemie.
- Piaget, J. (1948). *Psychologie der Intelligenz*. Rascher.
- Porsch, R., Baumgarten, M. & Jahn, R. W. (2023). Seiteneinsteiger:innen und traditionell ausgebildete Lehrkräfte: Unterschiede in den professionellen Kompetenzen, im beruflichen Handeln und emotionalen Erleben. In D. Behrens, M. Forell, T.-S. Idel & S. Pauling (Hrsg.), *Studien zur Professionsforschung und Lehrer:innenbildung. Lehrkräftebildung in der Bedarfskrise: Programme - Positionierungen - Empirie*. Verlag Julius Klinkhardt.
- Porst, R. (2011). Arten von Skalen. In R. Porst (Hrsg.), *Fragebogen* (S. 69–94). VS Verlag für Sozialwissenschaften.
https://doi.org/10.1007/978-3-531-92884-5_6
- Priyasmika, R. (2021). The Effect of Multiple Representation-Based Guided Inquiry on Learning Outcomes Reviewed from Scientific Thinking Skills. *EduChemia (Jurnal Kimia dan Pendidikan)*, 6(1), 55–66.
<https://doi.org/10.30870/educhemia.v6i1.8985>

- Reusser, K. & Pauli, C. (2014). Berufsbezogene Überzeugungen von Lehrerinnen und Lehrern. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 642–661). Waxmann.
- Richter, D. (2011). Lernen im Beruf. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 317–325). Waxmann.
- Roth, W.-M. & Pozzer-Ardenghi, L. (2013). Pictures in Biology Education. In D. F. Treagust & C.-Y. Tsui (Hrsg.), *Models and Modeling in Science Education: Bd. 7. Multiple Representations in Biological Education* (S. 39–54). Springer Netherlands.
- Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American psychologist*, 55(1), 68–78.
<https://doi.org/10.1037/0003-066X.55.1.68>
- Santos, V. C. & Arroio, A. (2016). The representational levels: Influences and contributions to research in chemical education. *Journal of Turkish Science Education*, 13(1), 3–18.
<https://doi.org/10.12973/tused.10153a>
- Scheeler, M. C., Ruhl, K. L. & McAfee, J. K. (2004). Providing Performance Feedback to Teachers: A Review. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, 27(4), 396–407.
<https://doi.org/10.1177/088840640402700407>
- Scheiter, K. & Richter, J. (2015). Multimediale Unterrichtsmaterialien gestalten: Ergebnisse der empirischen Lehr-Lernforschung. *Naturwissenschaften im Unterricht Chemie (Digitale Werkzeuge)*, 26(145), 8–11.
- Schmiemann, P. (2013). Aufgaben. *Unterricht Biologie*, 37(387/388), 2–8.
- Schnotz, W. (2002a). Commentary: Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, 14(1), 101–120.
<https://doi.org/10.1023/A:1013136727916>

- Schnotz, W. (2002b). Wissenserwerb mit Texten, Bildern und Diagrammen. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia und Internet: Lehrbuch für Studium und Praxis* (3., vollst. überarb. Aufl., S. 65–81). Beltz PVU.
- Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (2. Aufl., S. 72–103). Cambridge University Press.
- Schnotz, W. & Bannert, M. (1999). Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Text- und Bildverstehen. *Experimental Psychology*, 46(3), 217–236.
<https://doi.org/10.1026//0949-3964.46.3.217>
- Schnotz, W., Hauck, G. & Schwartz, N. H. (2022). Multiple mental representations in picture processing. *Psychological research*, 86(3), 903–918.
<https://doi.org/10.1007/s00426-021-01541-2>
- Schwarz, J. (2023). *Datenanalyse mit SPSS*. Universität Zürich.
https://www.methodenberatung.uzh.ch/de/datenanalyse_spss.html
- Sengutta, M. (2021). *Untersuchung der Qualität von externen multiplen Repräsentationen auf Arbeitsblättern von angehenden Lehrpersonen im Fach Chemie*. Masterarbeit. Universität Duisburg-Essen, Didaktik der Chemie.
- Seufert, T. (2003). Supporting coherence formation in learning from multiple representations: A Theoretical Framework and First Experimental Results. *Learning and Instruction*, 13(2), 227–237.
[https://doi.org/10.1016/S0959-4752\(02\)00022-1](https://doi.org/10.1016/S0959-4752(02)00022-1)
- Shapiro, S. S. & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4), 591.
<https://doi.org/10.2307/2333709>
- Shulman, L. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, 57(1), 1–23.
<https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Sliwka, H.-R. (2003). Reform of Chemical Language as a Model for Spelling Reform. *Journal of the Simplified Spelling Society*, 32, 24–28.
- Somers, R. H. (1962). A New Asymmetric Measure of Association for Ordinal Variables. *American Sociological Review*, 27(6), 799.
<https://doi.org/10.2307/2090408>

- Strietholt, R. & Terhart, E. (2009). Referendare beurteilen: Eine explorative Analyse von Beurteilungsinstrumenten in der Zweiten Phase der Lehrerbildung. *Zeitschrift für Pädagogik*, 55(4), 622–645.
<https://doi.org/10.25656/01:4266> (Zeitschrift für Pädagogik).
- Strübe, M. (2021). *Modelle und Experimente im Chemieunterricht* [Dissertation, Universität Duisburg-Essen; Logos Verlag Berlin]. GBV Gemeinsamer Bibliotheksverbund.
- Sweller, J., Ayres, P. & Kalyuga, S. (2011). *Cognitive load theory. Explorations in the learning sciences, instructional systems and performance technologies: [1]*. Springer.
<http://bvbr.bib-bvb.de:8991/F?func=service&doc%5Flibrary=BVB01&doc%5Fnumber=021195810&line%5Fnumber=0002&func%5Fcode=DB%5FRECORDS&service%5Ftype=MEDIA>
- Taber, K. S. (2009). Learning at the Symbolic Level. In J. K. Gilbert & D. F. Treagust (Hrsg.), *Multiple Representations in Chemical Education* (S. 75–105). Springer Netherlands.
- Talanquer, V. (2011). Macro, Submicro, and Symbolic: The many faces of the chemistry “triplet”. *International Journal of Science Education*, 33(2), 179–195.
<https://doi.org/10.1080/09500690903386435>
- Tan, K. C. D., Goh, N.-K., Chia, L. S. & Treagust, D. F. (2009). Linking the Macroscopic, Sub-microscopic and Symbolic Levels: The Case of Inorganic Qualitative Analysis. In J. K. Gilbert & D. F. Treagust (Hrsg.), *Multiple Representations in Chemical Education* (S. 137–150). Springer Netherlands.
- Taskin, V., Bernholt, S. & Parchmann, I. (2015a). An inventory for measuring student teachers’ knowledge of chemical representations: design, validation, and psychometric analysis. *Chem. Educ. Res. Pract.*, 16(3), 460–477.
<https://doi.org/10.1039/C4RP00214H>
- Taskin, V., Bernholt, S. & Parchmann, I. (2015b). Student Teachers’ Knowledge About Chemical Representations. *International Journal of Science and Mathematics Education*, 15(1), 39–55.
<https://doi.org/10.1007/s10763-015-9672-z>

- Taskin, V. & Bernholt, S. (2012). Students' Understanding of Chemical Formulae: A review of empirical research. *International Journal of Science Education*, 36(1), 157–185.
<https://doi.org/10.1080/09500693.2012.744492>
- Taskin, V., Bernholt, S. & Parchmann, I. (2012). Repräsentationsformen in der Chemie. In S. Bernholt (Hrsg.), *Gesellschaft für Didaktik der Chemie und Physik: Bd. 32. Konzepte fachdidaktischer Strukturierung für den Unterricht: Gesellschaft für Didaktik der Chemie und Physik; Jahrestagung in Oldenburg 2011* (S. 613–615). LIT.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S., Leutner, D., Neuhaus, B. J., Sandmann, A., Sumfleth, E., Thillmann, H. & Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7–28.
- Tima, M. T. & Sutrisno, H. (2018). Influence of Problem Solving Based on Multiple Representations Model on Teaching and Learning of Chemistry on Student's Academic Self-efficacy and Student's Cognitive Achievement. *American Journal of Educational Research*, 6(7), 887–892.
<https://doi.org/10.12691/education-6-7-1>
- Tippett, C. D. (2016). What recent research on diagrams suggests about learning with rather than learning from visual representations in science. *International Journal of Science Education*, 38(5), 725–746.
<https://doi.org/10.1080/09500693.2016.1158435>
- Tonyali, B. (2018). *Entwicklung und Evaluation eines Bewertungsbogens zur Analyse von Repräsentationen in selbsterstellten Lehr-Lern-Materialien von angehenden Lehrpersonen für den naturwissenschaftlichen Unterricht*. Masterarbeit. Universität Duisburg-Essen, Didaktik der Chemie.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2023). What makes representations good representations for science education? A teacher-oriented summary of significant findings and a practical guideline for the transfer into teaching. *Chemistry Teacher International*.
<https://doi.org/10.1515/cti-2022-0019>

- Treagust, D. F. (2018). The Importance of Multiple Representations for Teaching and Learning Science. In M. Shelley & S. A. Kiray (Hrsg.), *Education Research Highlights in Mathematics, Science and Technology 2018* (S. 215–223). International Society for Research in Education and Science (ISRES) Publishing.
- Treagust, D. F. & Chandrasegaran, A. L. (2009). The Efficacy of an Alternative Instructional Programme Designed to enhance Secondary Students' competence in the Triplet Relationship. In J. K. Gilbert & D. F. Treagust (Hrsg.), *Multiple Representations in Chemical Education* (S. 151–168). Springer Netherlands.
- Treagust, D. F., Duit, R. & Fischer, H. E. (Hrsg.). (2017). *Multiple Representations in Physics Education* (Bd. 10). Springer.
<https://doi.org/10.1007/978-3-319-58914-5>
- Treagust, D. F. & Tsui, C.-Y. (2013a). Foreword. In D. F. Treagust & C.-Y. Tsui (Hrsg.), *Models and Modeling in Science Education: Bd. 7. Multiple Representations in Biological Education* (S. V–VI). Springer Netherlands.
- Treagust, D. F. & Tsui, C.-Y. (Hrsg.). (2013b). *Models and Modeling in Science Education: Bd. 7. Multiple Representations in Biological Education*. Springer Netherlands.
- Tsui, C.-Y. & Treagust, D. F. (2013). Introduction to Multiple Repräsentations: Their Importance in Biology and Biological Education. In D. F. Treagust & C.-Y. Tsui (Hrsg.), *Models and Modeling in Science Education: Bd. 7. Multiple Representations in Biological Education* (S. 3–18). Springer Netherlands.
- Ude, S. (2016). *Entwicklung und Erprobung eines Kategoriensystems zur Analyse von Repräsentationen in Lernmaterialien für den naturwissenschaftlichen Unterricht*. Masterarbeit. Christian-Albrechts-Universität zu Kiel.
- van Driel, J. H. & Verloop, N. (1999). Teachers' knowledge of models and modelling in science. *International Journal of Science Education*, 21(11), 1141–1153.
<https://doi.org/10.1080/095006999290110>

- Walm, M. & Wittek, D. (2023). Lehrer:innenbildung in Deutschland – Überlegungen zur Situation und Perspektive der ‚Qualitätsfrage‘ in herausfordernden Zeiten. In D. Behrens, M. Forell, T.-S. Idel & S. Pauling (Hrsg.), *Studien zur Professionsforschung und Lehrer:innenbildung. Lehrkräftebildung in der Bedarfskrise: Programme - Positionierungen - Empirie* (S. 54–65). Verlag Julius Klinkhardt.
- Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S. & Garet, M. S. (2008). Experimenting With Teacher Professional Development: Motives and Methods. *Educational Researcher*, 37(8), 469–479.
<https://doi.org/10.3102/0013189X08327154>
- Weidenmann, B. (1988). *Psychische Prozesse beim Verstehen von Bildern* (1. Aufl.). *Huber-Psychologie-Forschung*. Huber.
- Widarti, H. R., Marfu’ah, S. & Parlan (2019). The Effects of Using Multiple Representations on Prospective Teachers’ Conceptual Understanding of Intermolecular Forces. *Journal of Physics: Conference Series*, 1227(1), 12006.
<https://doi.org/10.1088/1742-6596/1227/1/012006>
- Widarti, H. R., Permanasari, A., Mulyani, S., Ainur Rokhim, D. & Habiddin, H. (2021). Multiple Representation-Based Learning through Cognitive Dissonance Strategy to Reduce Student’s Misconceptions in Volumetric Analysis. *TEM Journal*, 1263–1273.
<https://doi.org/10.18421/TEM103-33>
- Wilcoxon, C. & Lemke, J. (2021). Preservice teachers’ perceptions of feedback: The importance of timing, purpose, and delivery. *Journal of University Teaching and Learning Practice*, 18(8).
<https://doi.org/10.53761/1.18.8.14>
- Wisniewski, B., Zierer, K. & Hattie, J. (2019). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in psychology*, 10.
<https://doi.org/10.3389/fpsyg.2019.03087>
- Wittrock, M. C. (1989). Generative Processes of Comprehension. *Educational Psychologist*, 24(4), 345–376.
https://doi.org/10.1207/s15326985ep2404_2

- Wiyarsi, A., Sutrisno, H. & Rohaeti, E. (2018). The effect of multiple representation approach on students' creative thinking skills: A case of 'Rate of Reaction' topic. *Journal of Physics: Conference Series*, 1097, 12054.
<https://doi.org/10.1088/1742-6596/1097/1/012054>
- Wollenschläger, M., Möller, J. & Harms, U. (2012). Ist kompetenzielles Feedback überlegen, weil es als effektiver wahrgenommen wird? *Unterrichtswissenschaft*, 40(3).
- Yore, L. D. & Hand, B. (2010). Epilogue: Plotting a Research Agenda for Multiple Representations, Multiple Modality, and Multimodal Representational Competency. *Research in Science Education*, 40(1), 93–101.
<https://doi.org/10.1007/s11165-009-9160-y>

7.2 Abkürzungsverzeichnis

*	Signifikanzniveau von $p < .050$ (signifikant)
**	Signifikanzniveau von $p < .005$ (hochsignifikant)
***	Signifikanzniveau von $p < .001$ (höchstsignifikant)
95% CI	Konfidenzintervall (engl. confidence interval) zum Niveau 95%
ANOVA	Analysis of Variance
B	Regressionskoeffizient
CL	Chemielehrkraft
CU	Chemieunterricht
DF	Darstellungsformen
ER	Externe Repräsentation
F	Maß der ANOVA für die Größe der Varianz
Fachdid.	fachdidaktisch
FDW	Fachdidaktisches Wissen
FW	Fachwissen
GymGe	Gymnasium und Gesamtschule (Lehramtsform)
H	Hauptschule
HRSGe	Haupt-, Real-, Sekundar- und Gesamtschule (Lehramtsform)
IG1	Interventionsgruppe 1 (Intervention mit internem und externem Feedback)
IG2	Interventionsgruppe 2 (Intervention mit internem Feedback)
KG	Kontrollgruppe (Intervention mit Dummy-Feedback)
Korr.	Korrigiert
LAA	Lehramtsanwärter:in (im Vorbereitungsdienst, abgeschlossenes Studium)
LiA	Lehrkraft in Ausbildung (Seiten-/Quereinstieg)
LLM	Lehr-Lern-Material
M	Mittelwert bzw. arithmetisches Mittel (engl. mean)
M_{Diff}	Mittelwertsdifferenz
MER	Multiple externe Repräsentation
MSB NRW	Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen
MZP	Messzeitpunkt
N	Anzahl

NRW	Nordrhein-Westfalen
OG	Obergrenze des Konfidenzintervalls
r	Korrelationskoeffizient
R	Realschule
R^2	Varianzaufklärung
Rep.	Repräsentationsbezogen
Res.	Residual
S	Schüler:innen
<i>SD</i>	Standardabweichung (engl. standard deviation)
SE	Standardfehler (engl. standard error)
Sek	Sekundarschule
SWE	Selbstwirksamkeitserwartung
t	Maß des t-Tests für die Größe der Varianz
U	Maß des Mann-Whitney U-Tests für die Größe der Varianz
UG	Untergrenze des Konfidenzintervalls
VD	Vorbereitungsdienst
ZfsL	Zentrum für schulpraktische Lehrerbildung
z-stand.	z-standardisiert
α	Cronbachs Alpha
β	Beta-Koeffizient
η^2	Effektstärke Eta-Quadrat einer ANOVA
χ^2	Maß des Chi-Quadrat-Tests für die Größe der Varianz

7.3 Tabellenverzeichnis

Tabelle 1:	Definition repräsentationsbezogener Kompetenzen von Lehrkräften (verändert nach Nitz, 2012)	21
Tabelle 2:	Klassifikation unterschiedlicher Feedbackformen (nach Narciss, 2006, S. 23).....	34
Tabelle 3:	Bedingungen der Wirksamkeit von Feedback zur Verbesserung des Unterrichtens (Eigene Zusammenstellung nach Brinko, 1993)	37
Tabelle 4:	Untersuchungsvariablen der Interventionsstudie	51
Tabelle 5:	Interventionsdesign.....	53
Tabelle 6:	Ablauf und Struktur einer Interventionssitzung am Beispiel der ersten Sitzung.....	56
Tabelle 7:	Inhaltliche Verteilung der Testitems zum allgemeinen Fachwissen.....	66
Tabelle 8:	Inhaltliche Verteilung der Testitems zum allgemeinen fachdidaktischen Wissen	67
Tabelle 9:	Inhaltliche Verteilung der Testitems zum repräsentationsbezogenen Fachwissen und repräsentationsbezogenen fachdidaktischen Wissen vor der Pilotierung und zugehörige Kürzungen.....	70
Tabelle 10:	Testkennwerte aus der Pilotierung zum repräsentationsbezogenen Fachwissen und repräsentationsbezogenen fachdidaktischen Wissen	71
Tabelle 11:	Inhaltliche Verteilung der Testitems zum repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen nach der Pilotierung und zugehörige Kürzungen	73
Tabelle 12:	Testkennwerte aus der Pilotierung zu den repräsentationsbezogenen Überzeugungen.....	74
Tabelle 13:	Vergleich der Ersatztätigkeiten der Untersuchungsgruppen für das interne und externe Feedback	80
Tabelle 14:	Bewertungsnormen für den Reliabilitätskoeffizienten Goodman und Kruskals Gamma (nach Döring & Bortz, 2023)...	82
Tabelle 15:	Übersicht zur Einordnung von Effektstärken bei Varianzanalysen und Korrelationen	86

Tabelle 16: Zur Auswertung ausgewählte Bewertungsnormen für den Reliabilitätskoeffizienten Cronbachs Alpha.....	87
Tabelle 17: Bewertungsnorm für R^2 zur Bewertung der Varianzaufklärung eines Regressionsmodells	89
Tabelle 18: Bewertungsnorm für R^2 zur Bewertung der Varianzaufklärung eines Regressionsmodells	89
Tabelle 19: Stichprobenzusammensetzung.....	92
Tabelle 20: Testkennwerte zu den persönlichen Ressourcen	94
Tabelle 21: Vergleich der mittleren Belastungen und Bewältigungsstrategien.....	95
Tabelle 22: Vergleich des mittleren professionellen Selbstverständnisses	95
Tabelle 23: Vergleich der mittleren individuellen beruflichen Selbstwirksamkeitserwartung	96
Tabelle 24: Testkennwerte zu den Erfahrungen in der Gestaltung von Lehr-Lern-Materialien	98
Tabelle 25: Vergleich der mittleren Erfahrungen in der Gestaltung von Lehr-Lern-Material	99
Tabelle 26: Testkennwerte zum allgemeinen Fachwissen und fachdidaktischen Wissen	100
Tabelle 27: Vergleich des mittleren relativen allgemeinen Fachwissens und fachdidaktischen Wissens.....	100
Tabelle 28: Testkennwerte zu den repräsentationsbezogenen Variablen.....	103
Tabelle 29: Vergleich des mittleren relativen repräsentationsbezogenen Fachwissens und fachdidaktischen Wissens	104
Tabelle 30: Vergleich der mittleren residualen Veränderung des repräsentationsbezogenen Fachwissens und fachdidaktischen Wissens	107
Tabelle 31: Vergleich der mittleren repräsentationsbezogenen Überzeugungen	109
Tabelle 32: Vergleich der residualen Veränderung der repräsentationsbezogenen Überzeugungen.....	110
Tabelle 33: Profillinien zu den Überzeugungen der Gesamtstichprobe im Prä-Post-Vergleich	114
Tabelle 34: Vergleich der mittleren Selbstbewertungsfähigkeit	124

Tabelle 35: Vergleich der mittleren Selbstbewertungsfähigkeit anhand einer dreistufigen Likert-Skala	128
Tabelle 36: Testkennwerte zum Cognitive Load und zur Anstrengungsbereitschaft	129
Tabelle 37: Vergleich des mittleren Cognitive Loads	131
Tabelle 38: Vergleich der mittleren Anstrengungsbereitschaft	132
Tabelle 39: Signifikante Varianzanalysen zwischen angehenden Lehrkräften unterschiedlicher Lehramtsformen	134
Tabelle 40: Signifikante Varianzanalysen zwischen angehenden Lehrkräften unterschiedlicher Hochschulabschlüsse	137
Tabelle 41: Signifikante Varianzanalysen zwischen angehenden Lehrkräften mit unterschiedlicher Dauer des bisher absolvierten bedarfsdeckenden Unterrichts	137
Tabelle 42: Signifikante Varianzanalysen zwischen angehenden Lehrkräften mit und ohne einer Vertretungsstelle vor dem Vorbereitungsdienst	138
Tabelle 43: Spearman-Korrelationen innerhalb der persönlichen Ressourcen	139
Tabelle 44: Pearson-Korrelationen zwischen den untersuchten Wissensbereichen im Prä- und Posttest	140
Tabelle 45: Spearman-Korrelationen zwischen den persönlichen Ressourcen und allgemeinen Wissensbereichen	141
Tabelle 46: Spearman-Korrelationen zwischen den persönlichen Ressourcen und den repräsentationsbezogenen Wissensbereichen	142
Tabelle 47: Pearson-Korrelationen zwischen den Wissensbereichen sowie persönlichen Ressourcen und den Qualitäten der selbstdesignten Lehr-Lern-Materialien	142
Tabelle 48: Spearman-Korrelationen zwischen den repräsentationsbezogenen Überzeugungen und dem professionellen Selbstverständnis	144
Tabelle 49: Pearson-Korrelationen zwischen dem Cognitive Load und relevanten Variablen	145

Tabelle 50: Exemplarisch ausgefüllter Bewertungsbogen und zugehörige Items zur Erfassung der Qualität der in der Intervention designten Lehr-Lern-Materialien	212
Tabelle 51: Itemkennwerte zum Test des rep. Fachwissens in der Pilotierung.....	221
Tabelle 52: Itemkennwerte zum Test des rep. fachdidaktischen Wissen in der Pilotierung	222
Tabelle 53: Itemkennwerte zum Fragebogen der rep. Überzeugungen in der Pilotierung	223
Tabelle 54: Itemkennwerte zum Test des rep. Fachwissens im Posttest	224
Tabelle 55: Itemkennwerte zum Test des rep. fachdidaktischen Wissens im Posttest.....	224
Tabelle 56: Itemkennwerte zum Fragebogen der rep. Überzeugungen im Posttest	225
Tabelle 57: Ergebnisse der gruppenspezifischen Varianzanalysen der demografischen Daten	226
Tabelle 58: Ergebnisse der signifikanten dreifaktoriellen Varianzanalysen zu den Gesamtmittelwerten der Interventionsgruppe 1 aus allen drei Lehr-Lern-Materialien (Nr. 1, 2 und 3).....	226
Tabelle 59: Ergebnisse der signifikanten dreifaktoriellen Varianzanalysen zu den Gesamtmittelwerten der Interventionsgruppe 2 aus allen drei Lehr-Lern-Materialien (Nr. 1, 2 und 3).....	226
Tabelle 60: Ergebnisse der signifikanten dreifaktoriellen Varianzanalysen zu den Gesamtmittelwerten der Kontrollgruppe aus allen drei Lehr-Lern-Materialien (Nr. 1, 2 und 3)	227
Tabelle 61: Prädiktoren des allgemeinen Fachwissens (prä) für die Gesamtstichprobe.....	227
Tabelle 62: Prädiktoren des allgemeinen fachdidaktische Wissens (prä) in der Gesamtstichprobe	228
Tabelle 63: Prädiktoren für das repräsentationsbezogene Fachwissen (prä) in der Gesamtstichprobe	228

Tabelle 64: Prädiktoren für das repräsentationsbezogene Fachwissen (post) in der Gesamtstichprobe	229
Tabelle 65: Prädiktoren für das repräsentationsbezogene fachdidaktische Wissen (prä) in der Gesamtstichprobe.....	229
Tabelle 66: Prädiktoren für das repräsentationsbezogene fachdidaktische Wissen (post) aufgeteilt nach den Untersuchungsgruppen	230
Tabelle 67: Prädiktoren für die repräsentationsbezogenen Überzeugungen (prä) in der Gesamtstichprobe	230
Tabelle 68: Prädiktoren für die Qualität der designten Lehr-Lern- Materialien (Nr. 1) in der Gesamtstichprobe.....	231

7.4 Abbildungsverzeichnis

Abbildung 1: Modell des integrativen Text- und Bildverstehens (Schnotz, 2014; Schnotz et al., 2022)	6
Abbildung 2: Funktionale Taxonomie multipler externer Repräsentationen (Ainsworth, 1999; übersetzt nach Krey & Schwanewedel, 2018)	8
Abbildung 3: Repräsentationsformen in der Chemie mit ausgewählten Beispielen (erstellt nach Gilbert & Treagust, 2009b; Lachmayer et al., 2007; Leisen, 2005; Mangold, 2015; Nerdel, 2017; Peterson, 1996; Treagust et al., 2017; Treagust & Tsui, 2013b; Weidenmann, 1988)	10
Abbildung 4: Abstraktionsgrade von externen Repräsentationen in den Naturwissenschaften am Beispiel des Blauen Lotus (verändert nach Roth & Pozzer-Ardenghi, 2013).....	11
Abbildung 5: Interactive Two Feedback-Loops Model (ITFL-Modell; aus Narciss, 2018).....	30
Abbildung 6: Testmodell zur Konstruktion der Testitems zum repräsentationsbezogenen Fachwissen und fachdidaktischen Wissen	69
Abbildung 7: Entscheidungspfad zur Auswahl der passenden Varianzanalysen für den vorliegenden Datensatz mit Beispielergebnissen	85
Abbildung 8: Absolvierte Studiengänge der angehenden Lehrkräfte	93
Abbildung 9: Relative Häufigkeiten zum Lernkontext der Gestaltung von Lehr-Lern-Materialien	98
Abbildung 10: Vergleich des mittleren relativen allgemeinen Fachwissens..	101
Abbildung 11: Vergleich des mittleren relativen allgemeinen fachdidaktischen Wissens.....	101
Abbildung 12: Vergleich des mittleren relativen repräsentationsbezogenen Fachwissens.....	105
Abbildung 13: Vergleich des mittleren relativen repräsentationsbezogenen fachdidaktischen Wissens.....	105
Abbildung 14: Vergleich der mittleren residualen Veränderung des repräsentationsbezogenen Fachwissens.....	107

Abbildung 15: Vergleich der mittleren residualen Veränderung des repräsentationsbezogenen fachdidaktischen Wissens	108
Abbildung 16: Vergleich der mittleren repräsentationsbezogenen Überzeugungen	109
Abbildung 17: Residuale Veränderungen bei den repräsentationsbezogenen Überzeugungen	110
Abbildung 18: Histogramm der mittleren repräsentationsbezogenen Überzeugungen der Gesamtgruppe	111
Abbildung 19: Mittlere Gesamtqualität der Lehr-Lern-Materialien der Gesamtgruppe	117
Abbildung 20: Mittlere Qualität des Lehr-Lern-Materialien Nr. 1 der Gesamtgruppe unterteilt nach den Qualitätsmerkmalen des Bewertungsbogens	118
Abbildung 21: Mittlere Qualität der Lehr-Lern-Materialien Nr. 2 der Gesamtgruppe unterteilt nach den Qualitätsmerkmalen des Bewertungsbogens	118
Abbildung 22: Mittlere Qualität des Lehr-Lern-Materials Nr. 3 für die Gesamtgruppe unterteilt nach den Qualitätsmerkmalen des Bewertungsbogens	119
Abbildung 23: Mittlere Qualität der Lehr-Lern-Materialien (Nr. 1, 2 und 3 zusammengefasst) der Gesamtgruppe unterteilt nach den Qualitätsmerkmalen des Bewertungsbogens	119
Abbildung 24: Mittlere Qualität der Lehr-Lern-Materialien der Interventionsgruppe 1 unterteilt nach den Sitzungen (Nr. 1, 2 und 3) und Qualitätsmerkmalen	121
Abbildung 25: Mittlere Qualität der Lehr-Lern-Materialien der Interventionsgruppe 2 unterteilt nach den Sitzungen (Nr. 1, 2 und 3) und Qualitätsmerkmalen	121
Abbildung 26: Mittlere Qualität der Lehr-Lern-Materialien der Kontrollgruppe unterteilt nach den Sitzungen (Nr. 1, 2 und 3) und Qualitätsmerkmalen	122
Abbildung 27: Mittlere Selbstbewertungsfähigkeit der Interventionsgruppen 1 und 2 unterteilt nach den Sitzungen (Nr. 1, 2, 3) ..	124

Abbildung 28: Mittlere Bewertungstendenz der Interventionsgruppe 1 bei ihren Lehr-Lern-Materialien unterteilt nach den Sitzungen (Nr. 1, 2 und 3) und Qualitätsmerkmalen.....	125
Abbildung 29: Bewertungstendenz der Interventionsgruppe 2 bei ihren Lehr-Lern-Materialien unterteilt nach den Sitzungen (Nr. 1, 2 und 3) und Qualitätsmerkmalen.....	126
Abbildung 30: Mittlere Selbstbewertungsfähigkeit unterteilt nach den Sitzungen (Nr. 1, 2, 3) und jeweils ausgewertet mit einer fünf- und dreistufigen Likert-Skala.....	128
Abbildung 31: Histogramm des Cognitive Loads sowie der Anstrengungsbereitschaft in den Interventionsgruppen 1 und 2, unterteilt nach den Sitzungen	130
Abbildung 32: Mittlerer wahrgenommener Cognitive Load in den Interventionsgruppen 1 und 2 während der Feedbackphasen.....	131
Abbildung 33: Mittlere wahrgenommene Anstrengungsbereitschaft in den Interventionsgruppe 1 und 2 während der Feedbackphasen ..	132
Abbildung 34: Vergleich des mittleren relativen allgemeinen Fachwissens unterteilt nach Lehramtsformen.....	135
Abbildung 35: Vergleich des mittleren relativen repräsentationsbezogenen Fachwissens unterschiedlicher Lehramtsformen.....	135
Abbildung 36: Streuung zwischen dem allgemeinen Fachwissen im Prätest und repräsentationsbezogenen Fachwissen im Posttest	146
Abbildung 37: Streuung des repräsentationsbezogenen Fachwissens zwischen Prä- und Posttest	148
Abbildung 38: Streuung zwischen dem allgemeinen fachdidaktischen Wissen im Prätest und repräsentationsbezogenen fachdidaktischen Wissen im Posttest.....	149
Abbildung 39: Streuung des repräsentationsbezogenen fachdidaktischen Wissens zwischen Prä- und Posttest.....	150
Abbildung 40: Prädiktoren des allgemeinen Fachwissens und fachdidaktischen Wissens mit zugehörigen β -Koeffizienten	152

Abbildung 41: Prädiktoren des repräsentationsbezogenen Fachwissens im Prä- und Posttest mit zugehörigen β -Koeffizienten	152
Abbildung 42: Prädiktoren des repräsentationsbezogenen fach- didaktischen Wissens mit zugehörigen β -Koeffizienten, im Posttest unterteilt nach den Untersuchungsgruppen	153
Abbildung 43: Bildlicher Teil des Rohmaterials für die erste Interventionssitzung.....	205
Abbildung 44: Textlicher Teil des Rohmaterials für die erste Interventionssitzung.....	206
Abbildung 45: Bildlicher Teil des Rohmaterials für die zweite Interventionssitzung.....	207
Abbildung 46: Textlicher Teil des Rohmaterials für die zweite Interventionssitzung.....	207
Abbildung 47: Bildlicher Teil des Rohmaterials für die dritte Interventionssitzung.....	208
Abbildung 48: Textlicher Teil des Rohmaterials für die dritte Interventionssitzung (Thema: Stoffgemische anhand des einfachen Teilchenmodells).....	209

8 ANHANG

A Interventionsmaterialien und Testinstrumente

A.1 Einsicht in die Moodle-Kurse

Im Folgenden wird die Struktur der Moodle-Kurse am Beispiel der Interventionsgruppe 1 dargestellt. Es sind zum einen die Abfragemasken zum Prä- und Post-Test und zum anderen die Interventionsinhalte abgebildet. Die Interventionsinhalte enthalten Material sowohl für das interne als auch externe Feedback.

Vor Beginn der 1. Sitzung (Vortests)

Vor Beginn der eigentlichen drei Sitzungen möchten wir Ihr Vorwissen und Ihre Meinungen messen. Bitte nehmen Sie sich hierfür 105 Minuten Zeit, um die Fragen nach möglichst bestem Gewissen zu beantworten.

Die Zeitbegrenzungen der sieben Frageteile sind auf den unteren Links in Klammern gesetzt. Bitte halten Sie sich selbstständig an die zeitlichen Vorgaben und nutzen Sie für die Antworten keine Hilfsmittel. Nur so ist eine zuverlässige Erfassung Ihres Wissens möglich.

Wichtig: Damit der Test nach Abschluss gespeichert wird, klicken Sie erst auf „Versuch beenden...“ und zusätzlich auf den „**Abgabe**“-Button. Wenn Sie den „**Abgabe**“-Button nicht anklicken, wird der Test nicht von allein gespeichert und Ihre Ergebnisse gehen verloren.

-  1. Meinungsabfrage zum Chemieunterricht (15 min)
-  2. Fachdidaktisches Wissen zu Darstellungsformen (20 min)
-  3. Fachwissen zu Darstellungsformen (20 min)
-  4. Allgemeines fachdidaktisches Wissen (15 min)
-  5. Allgemeines Fachwissen (15 min)
-  6. Allgemeine Angaben I (10 min)
-  7. Allgemeine Angaben II (10 min)

Nach Abschluss der 3. Sitzung (Nachtests)

Nehmen Sie sich zum Abschluss 80 Minuten Zeit und bearbeiten Sie die unten aufgeführten Fragen.

Die Zeitbegrenzungen der fünf Frageteile sind auf den unteren Links in Klammern gesetzt. Bitte halten Sie sich selbstständig an die zeitlichen Vorgaben und nutzen Sie für die Antworten keine Hilfsmittel. Nur so ist eine zuverlässige Erfassung Ihres Wissens möglich.

Sobald Sie alle Vorher-Nachher-Befragungen absolviert und die drei Aufgaben der Sitzungen abgegeben haben, senden Sie uns an fire2@uni-due.de eine kurze E-Mail, damit wir Ihnen Ihre Probandenvergütung zukommen lassen.

Im Namen des gesamten FiRe2-Teams bedanken wir uns herzlich bei Ihnen für Ihre Zeit und Mühe!

Wichtig: Damit der Test nach Abschluss gespeichert wird, klicken Sie erst auf „Versuch beenden...“ und zusätzlich auf den „Abgabe“-Button. Wenn Sie den „Abgabe“-Button nicht anklicken, wird der Test nicht von allein gespeichert und Ihre Ergebnisse gehen verloren.

-  1. Meinungsabfrage zum Chemieunterricht (10 min)
-  2. Fachdidaktisches Wissen zu Darstellungsformen (20 min)
-  3. Fachwissen zu Darstellungsformen (20 min)
-  4. Allgemeines fachdidaktisches Wissen (15 min)
-  5. Allgemeines Fachwissen (15 min)

1. Sitzung

Herzlich Willkommen zu Ihrer ersten Sitzung im Lernmodul! In dieser Sitzung lernen Sie die Begrifflichkeiten zu Darstellungsformen kennen und gestalten daran anlehnend Ihr erstes Arbeitsblatt. Bitte gehen Sie bei der Bearbeitung der Sitzung anhand der folgenden Schritte vor:

Schritt 1:

Öffnen Sie die Präsentation "Einführung und Arbeitsauftrag 1" und lesen Sie die Folien aufmerksam durch.

Schritt 2:

Laden Sie die "Rohdatei Arbeitsblatt 1" auf Ihren Computer herunter. Beginnen Sie wie in der Präsentation beschrieben mit der Überarbeitung des Arbeitsblattes. Für die Anfertigung stehen Ihnen **45 Minuten** zur Verfügung.

Schritt 3:

Speichern Sie Ihr Arbeitsblatt als PDF-Dokument ab und laden Sie es in "Abgabe für Ihr fertiges Arbeitsblatt 1" hoch.

Schritt 4:

Führen Sie die Bewertung Ihres eigenen Arbeitsblattes mit dem "Bewertungsbogen für Ihr Arbeitsblatt 1" durch. Sie können hierbei jederzeit nochmals in Ihr Arbeitsblatt sehen. Für die Bewertung stehen Ihnen **30 Minuten** zur Verfügung.

Schritt 5:

Starten Sie zu Ihrem nächsten Sitzungstermin die **2. Sitzung**. Bis dahin fertigen wir unsere individuelle Rückmeldung für Ihr Arbeitsblatt an.

-  Einführung und Arbeitsauftrag 1
-  Rohdatei Arbeitsblatt 1 Teilchenmodell (Bearbeitungszeit: 45 min)
-  Abgabe für Ihr fertiges Arbeitsblatt 1
-  Bewertungsbogen für Ihr Arbeitsblatt 1 (Bearbeitungszeit: 30 min)

2. Sitzung

Willkommen zu Ihrer zweiten Sitzung! In dieser Sitzung sollen Sie Ihre neu gewonnen Erkenntnisse und Erfahrungen aus der ersten Sitzung reflektieren, anwenden und ausbessern. Bitte gehen Sie bei der Bearbeitung der Sitzung anhand der folgenden Schritte vor:

Schritt 6:

Öffnen Sie in dem unteren Ordner "**Rückmeldungen vom FiRe2-Team zu Ihrem Arbeitsblatt 1**" die individuelle Rückmeldung für Ihr Arbeitsblatt 1. Vergleichen Sie Ihre eigenen Bewertungen mit denen des FiRe₂-Teams. Notieren Sie sich ggf. in Ihren persönlichen Notizen Punkte oder Aspekte, die Sie in Ihrem zweiten Arbeitsblatt besser gestalten möchten. Ihnen stehen **15 Minuten** zur Verfügung.

Schritt 7:

Nun beginnt der bisherige Lernzyklus von neu. Öffnen Sie die Präsentation "**Arbeitsauftrag 2**" und lesen Sie die Folien aufmerksam durch.

Schritt 8:

Laden Sie die "**Rohdatei Arbeitsblatt 2**" auf Ihren Computer herunter. Beginnen Sie wie in der Präsentation beschrieben mit der Überarbeitung des Arbeitsblattes. Für die Anfertigung stehen Ihnen **45 Minuten** zur Verfügung.

Schritt 9:

Speichern Sie Ihr Arbeitsblatt als PDF-Dokument ab und laden Sie es in "**Abgabe für Ihr fertiges Arbeitsblatt 2**" hoch.

Schritt 10:

Führen Sie die Bewertung Ihres eigenen Arbeitsblattes mit dem "**Bewertungsbogen für Ihr Arbeitsblatt 2**" durch. Sie können hierbei jederzeit nochmals in Ihr Arbeitsblatt sehen. Für die Bewertung stehen Ihnen **20 Minuten** zur Verfügung.

Schritt 11:

Starten Sie zu Ihrem nächsten Sitzungstermin die **3. Sitzung**. Bis dahin fertigen wir unsere individuelle Rückmeldung für Ihr Arbeitsblatt an.

-  Rückmeldungen vom Fire2-Team zu Ihrem Arbeitsblatt 1
-  Arbeitsauftrag 2
-  Rohdatei Arbeitsblatt 2 Aggregatzustände (Bearbeitungszeit: 45 min)
- 
-  Abgabe für Ihr fertiges Arbeitsblatt 2
-  Bewertungsbogen für Ihr Arbeitsblatt 2 (Bearbeitungszeit: 20 min)

3. Sitzung

Willkommen zu Ihrer dritten Sitzung! In dieser Sitzung sollen Sie Ihre neu gewonnen Erkenntnisse und Erfahrungen aus den vorigen Sitzungen nochmals anwenden und sichern. Bitte gehen Sie bei der Bearbeitung der Sitzung anhand der folgenden Schritte vor:

Schritt 12:

Öffnen Sie in dem unteren Ordner "**Rückmeldungen vom FiRe2-Team zu Ihrem Arbeitsblatt 2**" die individuelle Rückmeldung für Ihr Arbeitsblatt 2. Vergleichen Sie Ihre eigenen Bewertungen mit denen des FiRe₂-Teams. Notieren Sie sich ggf. in Ihren persönlichen Notizen Punkte oder Aspekte, die Sie in Ihrem zweiten Arbeitsblatt besser gestalten möchten. Ihnen stehen **15 Minuten** zur Verfügung.

Schritt 13:

Nun beginnt der Lernzyklus nochmals von neu. Öffnen Sie die Präsentation "**Arbeitsauftrag 3**" und lesen Sie die Folien aufmerksam durch.

Schritt 14:

Laden Sie die "**Rohdatei Arbeitsblatt 3**" auf Ihren Computer herunter. Beginnen Sie wie in der Präsentation beschrieben mit der Überarbeitung des Arbeitsblattes. Für die Anfertigung stehen Ihnen **45 Minuten** zur Verfügung.

Schritt 15:

Speichern Sie Ihr Arbeitsblatt als PDF-Dokument ab und laden Sie es in "**Abgabe für Ihr fertiges Arbeitsblatt 3**" hoch.

Schritt 16:

Führen Sie die Bewertung Ihres eigenen Arbeitsblattes mit dem "**Bewertungsbogen für Ihr Arbeitsblatt 3**" durch. Sie können hierbei jederzeit nochmals in Ihr Arbeitsblatt sehen. Für die Bewertung stehen Ihnen **20 Minuten** zur Verfügung.

Schritt 17:

Nun haben Sie alle Sitzungen erfolgreich abgeschlossen! Bitte führen Sie zu Ihrem letzten Termin den Nachtest durch.

 Rückmeldungen vom Fire2-Team zu Ihrem Arbeitsblatt 2

 Arbeitsauftrag 3

 Rohdatei Arbeitsblatt 3 (Bearbeitungszeit: 45 min)



 Abgabe für Ihr fertiges Arbeitsblatt 3

 Bewertungsbogen für Ihr Arbeitsblatt 3 (Bearbeitungszeit: 20 min)

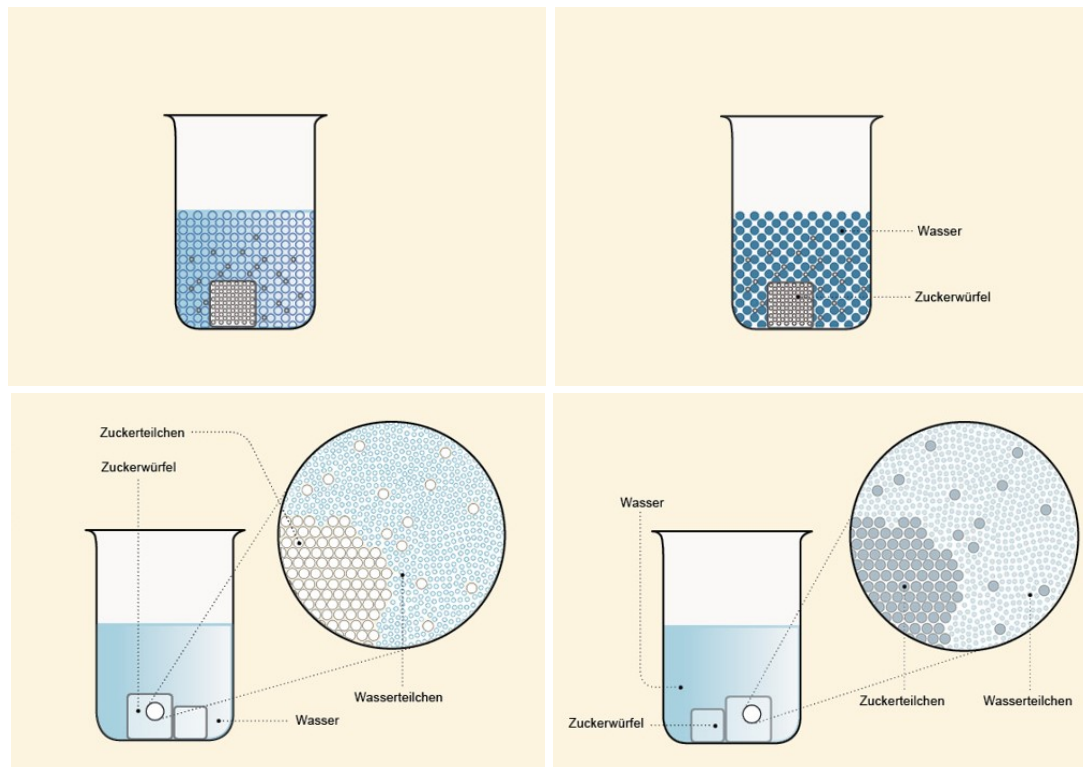
A.2 Bereitgestellte Bild- und Textelemente in den Rohmaterialien**Interventionssitzung 1: Aufbau von Materie anhand des einfachen Teilchenmodells**

Abbildung 43: Bildlicher Teil des Rohmaterials für die erste Interventionssitzung

Gibt man Zucker in den Tee, sind die Zuckerkristalle nach einer kurzen Zeit nicht mehr sichtbar. Eine Geschmacksprobe zeigt aber, dass der Tee süß schmeckt. Der Zucker ist also immer noch vorhanden und hat sich in dem Tee nur gelöst.

Das Teilchenmodell

Dieses Phänomen lässt sich erklären, wenn man sich vorstellt, dass Stoffe aus sehr, sehr kleinen, nicht sichtbaren Teilchen bestehen. Nach dem Teilchenmodell bestehen alle Stoffe kleinsten Teilchen, die sich ständig bewegen. Wir können sie uns als farbige Kugeln vorstellen. Die Atome eines einzigen Stoffes (z. B. Wasser) sind untereinander alle gleich. Sie sind gleich groß und gleich schwer. Die Teilchen verschiedener Stoffe (z. B. von Zucker und Wasser) unterscheiden sich voneinander.

Modelle zur Erklärung von Phänomenen

Zwischen dem Teilchenmodell und der Wirklichkeit gibt es jedoch Unterschiede. Modelle sind exakte Darstellungen und helfen uns, Phänomene

zu erklären, die wir mit dem bloßen Auge nicht verstehen können. Das Teilchenmodell beschreibt daher auch nicht, wie die Teilchen tatsächlich aussehen. Es erklärt aber bestimmte Sachverhalte und Vorgänge, so wie das Lösen von Zucker in Wasser.

Zusammenfassung des Teilchenmodells

Materie besteht aus kleinsten Teilchen. Im Raum zwischen den Teilchen befindet sich nichts

Die Teilchen verschiedener Stoffe unterscheiden sich in ihrer Größe und Masse.

Zwischen den Teilchen wirken Anziehungskräfte.

Die Teilchen sind ständig in Bewegung.

Mit sinkender Temperatur bewegen sich die Teilchen heftiger.

In Feststoffen haben die Teilchen eine dichte Anordnung und schwingen um ihre Plätze. In Flüssigkeiten bewegen sie sich ungeordnet, berühren sich aber gegenseitig. Bei Gasen sind die Anziehungskräfte nicht vorhanden. Die Abstände zwischen den Teilchen sind sehr groß.

Abbildung 44: Textlicher Teil des Rohmaterials für die erste Interventionssitzung

Interventionssitzung 2: Aggregatzustände anhand des einfachen Teilchenmodells

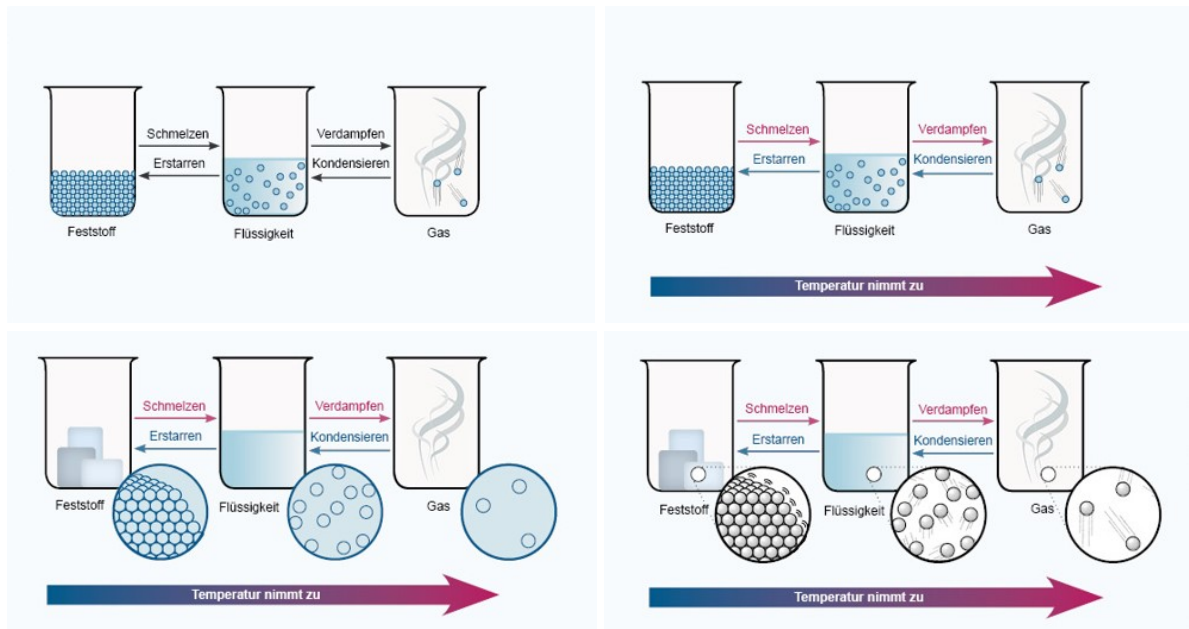


Abbildung 45: Bildlicher Teil des Rohmaterials für die zweite Interventionssitzung

Jeder Stoff kann in einem festen, flüssigen oder gasförmigen Zustand vorkommen. Diese drei Zustandsformen werden Aggregatzustände genannt. Wie die Aggregatzustände entstehen und wie ihre Übergänge ablaufen, können wir auf der Teilchenebene mithilfe des Teilchenmodells erklären.

In der Abbildung 2 ist ein Überblick über die Aggregatzustände zu sehen. Feststoff: Die Wasserteilchen im Eis sind regelmäßig angeordnet und bewegen sich kaum. Die Anziehungskräfte zwischen den Teilchen sorgen dafür, dass der Feststoff zusammenhält. Flüssigkeit: Beim Erwärmen wackeln die Teilchen hin und her und sie gehen voneinander weg. Schließlich so stark, dass die Teilchen ihre Plätze verlassen. Die Anordnung der Teilchen geht verloren. Der Feststoff schmilzt. Je größer die Anziehungskräfte zwischen den Teilchen sind, desto höher ist die Schmelztemperatur des Stoffes. Eis

wird wässrig bei 0°C. Gas: Wird das flüssige Wasser weiter erhitzt, so bewegen sich die Teilchen noch schneller. Die Abstände zwischen den Wasserteilchen werden größer, sodass die Anziehungskräfte kaum noch wirksam sind. Die Teilchen verlieren ihren Zusammenhalt: Das Wasser kocht.

Kühlt man gasförmigen Wasserdampf ab, so erstarrt das Wasser. Die Teilchenbewegung verringert sich, die Teilchen gehen näher aneinander. Es entsteht eine Flüssigkeit. Bei weiterem Abkühlen unter die Schmelztemperatur erstarrt das Wasser zu Eis. Die Bewegung der Teilchen wird noch geringer. Sie rücken näher zusammen und ordnen sich regelmäßig an. Diese regelmäßige Anordnung könnte man gut bei der Bildung von Eiskristallen beobachten.

Abbildung 46: Textlicher Teil des Rohmaterials für die zweite Interventionssitzung

Interventionssitzung 3: Stoffgemische anhand des einfachen Teilchenmodells

Heterogene Gemische

Feststoff und Feststoff



Granitgestein: ein Gemenge

Feststoff und Flüssigkeit



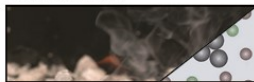
Organgensaft: eine Suspension

Flüssigkeit und Flüssigkeit



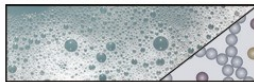
Milch: eine Emulsion

Feststoff und Gas



Grillfeuer: ein Rauch

Flüssigkeit und Gas



Badeschaum: ein Schaum

Homogene Gemische

Feststoff und Feststoff



Messing: eine Legierung

Feststoff und Flüssigkeit



Salzwasser: eine Lösung

Flüssigkeit und Flüssigkeit



Essig: eine Lösung

Feststoff und Gas



Wasserstoff-Platin-Katalysator: Lösung

Flüssigkeit und Gas



Sprudelwasser: eine Lösung

Heterogene Gemische

Gemenge = Feststoff + Feststoff



Beispiel: Granitgestein

Suspension = Feststoff + Flüssigkeit



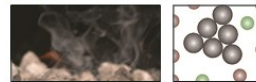
Beispiel: Organgensaft

Emulsion = Flüssigkeit + Flüssigkeit



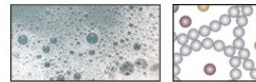
Beispiel: Milch

Rauch = Feststoff + Gas



Beispiel: Grillfeuer

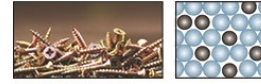
Schaum = Flüssigkeit + Gas



Beispiel: Badeschaum

Homogene Gemische

Legierung = Feststoff + Feststoff



Beispiel: Messing

Lösung = Feststoff + Flüssigkeit



Beispiel: Salzwasser

Lösung = Flüssigkeit + Flüssigkeit



Beispiel: Essig

Lösung = Feststoff + Gas



Beispiel: Wasserstoff-Platin-Katalysator

Lösung = Flüssigkeit + Gas



Beispiel: Sprudelwasser

Heterogene Gemische

Gemenge (Feststoff und Feststoff)

Beispiel: Granitgestein



Suspension (Feststoff und Flüssigkeit)

Beispiel: Organgensaft



Emulsion (Flüssigkeit und Flüssigkeit)

Beispiel: Milch



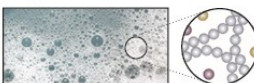
Rauch (Feststoff und Gas)

Beispiel: Grillfeuer



Schaum (Flüssigkeit und Gas)

Beispiel: Badeschaum



Homogene Gemische

Legierung (Feststoff und Feststoff)

Beispiel: Messing



Lösung (Feststoff und Flüssigkeit)

Beispiel: Salzwasser



Lösung (Flüssigkeit und Flüssigkeit)

Beispiel: Essig



Lösung (Feststoff und Gas)

Beispiel: Wasserstoff-Platin-Katalysator



Lösung (Flüssigkeit und Gas)

Beispiel: Sprudelwasser



Abbildung 47: Bildlicher Teil des Rohmaterials für die dritte Interventionssitzung

Ein Vorratsglas mit Müsli und ein Vorratsglas mit Zucker stehen nebeneinander. Deutlich kannst du im Müsli verschiedene Bestandteile unterscheiden, während der Zucker überall gleich aussieht.

Zucker besteht ausschließlich aus Zuckerkristallen. Alle Zuckerkristalle haben dieselben Stoffeigenschaften und lassen sich nicht voneinander unterscheiden. Stoffe wie Zucker, die nur aus einer einzigen Stoffart bestehen, werden Reinstoffe genannt. Müsli dagegen besteht aus mehreren Stoffen und ist daher ein Stoffgemisch. Im Haushalt findest du viele solcher Stoffgemische, wie zum Beispiel Waschmittel, Medikamente oder in Kosmetika, worin immer mehrere Stoffe miteinander vermischt sind.

Verschiedene Gemische

Bei einigen Stoffgemischen erkennt man die unterschiedlichen Bestandteile auf den ersten Blick, zum Beispiel bei einem Müsli. Bei anderen Stoffgemischen kann man einzelne Bestandteile noch mit einer Lupe oder einem Mikroskop voneinander unterscheiden, zum Beispiel bei Brausepulver. Solche Stoffgemische, deren unterschiedliche Bestandteile mit den Augen, einer Lupe oder dem Mikroskop zu erkennen sind, bezeichnet man als homogene Stoffgemische. Löst man etwas Kochsalz oder Zucker in

Wasser auf, dann sieht das entstandene Gemisch wie ein einheitlicher Stoff aus. Auch unter dem Mikroskop erkennt man die gelösten Bestandteile nicht mehr. Daher werden Stoffgemische, deren unterschiedliche Bestandteile nicht mehr erkennbar sind, heterogene Stoffgemische genannt.

Fachbegriffe für Stoffgemische

Wenn man wasserlösliche Stoffe wie Kochsalz oder Zucker mit Wasser verrührt, dann erhält man eine Lösung. Verrührt man dagegen einen wasserunlöslichen Stoff wie Kalk, Lehm oder Sand mit Wasser, so entsteht eine Suspension. Nach einiger Zeit setzen sich die festen Bestandteile dieses Gemisches am Boden ab. Bei einer Emulsion ist eine Flüssigkeit in feinen Tröpfchen in einer anderen Flüssigkeit verteilt. Zu den Emulsionen gehören beispielsweise Milch und Mayonnaise. Verteilen sich feinste Flüssigkeitströpfchen in einem Gas, dann entsteht ein Nebel. Wenn sich feinste Feststoffpartikel in einem Gas verteilen, wird das Gemisch als Rauch bezeichnet. Im Rauch eines Feuers vermischen sich feiner Ruß und Staub mit der Luft.

Reinstoffe sind Stoffe, die aus nur einem Stoff bestehen. Stoffgemische enthalten mindestens zwei Reinstoffe.

Abbildung 48: Textlicher Teil des Rohmaterials für die dritte Interventionssitzung (Thema: Stoffgemische anhand des einfachen Teilchenmodells)

A.3 Exemplarisches Word-Dokument des Rohmaterials für die erste Interventionssitzung

Im Folgenden ist das Rohmaterial aus der Sitzung mit dem Thema zur Einführung des einfachen Teilchenmodells zu finden. Das Rohmaterial zur zweiten und dritten Sitzung ist in identischer Form aufgebaut, jedoch für andere chemische Themen ausgelegt.

Die unteren Informationen auf Seite 1 sind für Sie als Notizen gedacht. Löschen Sie diesen Text, spätestens, wenn Sie mit der Gestaltung Ihres Arbeitsblattes fertig sind und ihre Datei hochladen werden.

Informationen aus den Präsentationsfolien zum Arbeitsauftrag

Arbeitsauftrag

Gestalten Sie aus dem bereitgestellten Dokument "Rohdatei Arbeitsblatt" ein Arbeitsblatt für Ihren Chemieunterricht. Laden Sie hierzu diese Datei herunter und nutzen Sie ein Textverarbeitungsprogramm (Microsoft Office, Apache OpenOffice o.ä.). Ihnen steht eine Bearbeitungszeit von 45 min zur Verfügung. Wenn Ihr Arbeitsblatt fertig ist, speichern Sie es als eine PDF-Datei ab und laden Sie es unter "Abgabe für Ihr fertiges Arbeitsblatt 1" hoch.

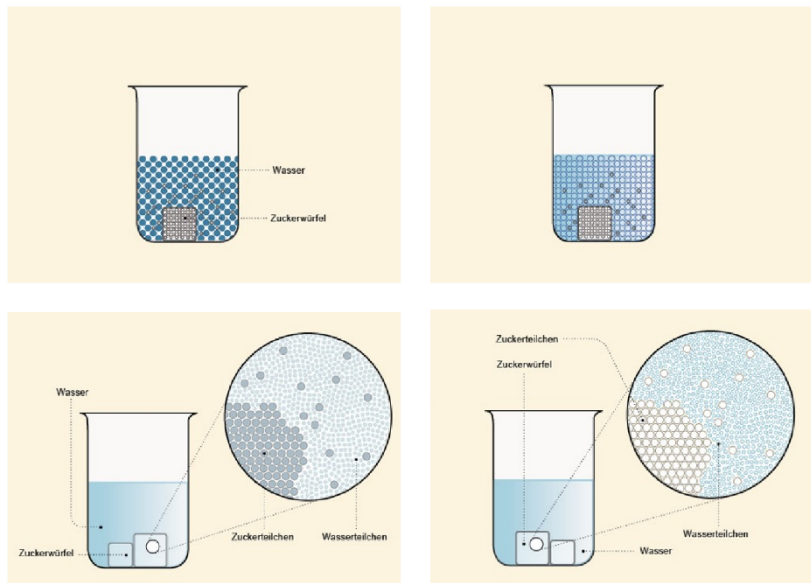
Beachten Sie bei Gestaltung des Arbeitsblattes (AB) folgende Regeln:

1. Halten Sie sich an die Vorgaben des fiktiven Lernziels und der Klasse (s.u.).
2. Ihr AB soll vollständig und selbsterklärend sein. Stellen Sie sich vor, dass Sie es zu Beginn der Unterrichtsstunde austeilen und jede/-r Schüler/-in das AB selbstständig und in Einzelarbeit durcharbeitet. Am Ende der Unterrichtsstunde sollte jeder/r Schüler/-in ohne fremde Hilfe das Lernziel erreicht haben.
3. In dem Dokument finden Sie eine Auswahl an Abbildungen. Suchen Sie eine dieser Abbildungen aus und nutzen Sie diese für Ihr AB. Sie können an Ihrer ausgewählten Abbildung auch zusätzliche Überarbeitungen mithilfe von anderen Programmen vornehmen. Bitte nutzen Sie aber keine anderen als diese gestellten Abbildungen für Ihr AB.
4. Ergänzen Sie Ihre ausgewählte und ggf. überarbeitete Abbildung mit einem geeigneten Text. Richten Sie sich dabei nach den Textfragmenten in dem Dokument. Den Text dürfen Sie vollständig sowie in Teilen verändern, überarbeiten, löschen, ergänzen, ersetzen und formatieren. Hier gibt es keine Begrenzungen.
5. Formulieren Sie geeignete Aufgabenstellungen für Ihre Schüler*innen.

Beachten Sie zudem folgende unterrichtliche Situation:

Sie befinden sich am Anfang des Schuljahres und unterrichten eine Klasse der **Jahrgangsstufe 7**. Ihre Schüler*innen haben dieses Jahr zum ersten Mal Chemie. Bisher haben Sie 6-7 Stunden unterrichtet. Inhaltlich beschäftigen Sie sich immer noch mit dem Inhaltsfeld "**Stoffe und Stoffeigenschaften (1)**". Die Klasse ist weitestgehend homogen, sodass Sie auf besondere Bedürfnisse einzelner Schüler*innen mit Förderbedarf *nicht* eingehen müssen. Die Schüler*innen bekommen in dieser Stunde für das Arbeitsblatt eine Bearbeitungszeit von 30 min. Das Lernziel bzw. die zu fördernde Kompetenz für das zu gestaltende Arbeitsblatt lautet:

Die Schüler*innen können den Aufbau von Stoffen und Materie am Beispiel des Kugelteilchenmodells beschreiben.



Gibt man Zucker in den Tee, sind die Zuckerkristalle nach einer kurzen Zeit nicht mehr sichtbar. Eine Geschmacksprobe zeigt aber, dass der Tee süß schmeckt. Der Zucker ist also immer noch vorhanden und hat sich in dem Tee nur gelöst.

Das Teilchenmodell

Dieses Phänomen lässt sich erklären, wenn man sich vorstellt, dass Stoffe aus sehr, sehr kleinen, nicht sichtbaren Teilchen bestehen. Nach dem Teilchenmodell bestehen alle Stoffe kleinsten Teilchen, die sich ständig bewegen. Wir können sie uns als farbige Kugeln vorstellen. Die Atome eines einzigen Stoffes (z.B. Wasser) sind untereinander alle gleich. Sie sind gleich groß und gleich schwer. Die Teilchen verschiedener Stoffe (z.B. von Zucker und Wasser) unterscheiden sich voneinander.

Modelle zur Erklärung von Phänomenen

Zwischen dem Teilchenmodell und der Wirklichkeit gibt es jedoch Unterschiede. Modelle sind exakte Darstellungen und helfen uns, Phänomene zu erklären, die wir mit dem bloßen Auge nicht verstehen können. Das Teilchenmodell beschreibt daher auch nicht, wie die Teilchen tatsächlich aussehen. Es erklärt aber bestimmte Sachverhalte und Vorgänge, so wie das Lösen von Zucker in Wasser.

Zusammenfassung des Teilchenmodells

Materie besteht aus kleinsten Teilchen. Im Raum zwischen den Teilchen befindet sich nichts

Die Teilchen verschiedener Stoffe unterscheiden sich in ihrer Größe und Masse.

Zwischen den Teilchen wirken Anziehungskräfte.

Die Teilchen sind ständig in Bewegung.

Mit sinkender Temperatur bewegen sich die Teilchen heftiger.

In Feststoffen haben die Teilchen eine dichte Anordnung und schwingen um ihre Plätze. In Flüssigkeiten bewegen sie sich ungeordnet, berühren sich aber gegenseitig. Bei Gasen sind die Anziehungskräfte nicht vorhanden. Die Abstände zwischen den Teilchen sind sehr groß.

A.4 Test zur Erfassung der repräsentationsbezogenen Qualität von Lehr-Lern-Materialien

Der Bewertungsbogen wurde als Interventionsmaterial und Testinstrument zugleich eingesetzt. Zugehörige Erläuterungen sind im Methodenteil zu finden.

Anmerkung zum Einsatz. Der Bewertungsbogen ist inhaltlich auf die Lernziele und Repräsentationen der vorliegenden Intervention reduziert und angepasst worden. Es besteht ein Zugriff auf eine universelle Handreichung im Paper-Pencil-Format (Mertner, 2021), welche auf Anfrage zur Verfügung gestellt werden kann. Der angepasste Bewertungsbogen ist zudem als Moodle-Tool übertragbar und verfügbar.

Tabelle 50: Exemplarisch ausgefüllter Bewertungsbogen und zugehörige Items zur Erfassung der Qualität der in der Intervention designten Lehr-Lern-Materialien

Bewertungsbogen für Lehr-Lern-Materialien		
Besitzer:in des Arbeitsblattes: IEL01		
Thema: (x) 1. Teilchenmodell, () 2. Aggregatzustände, () 3. Stoffgemische		
Themen und Kriterien	Eigene Bewertung ↓	FiRe2-Bewertung ↓
1. Übereinstimmung mit Lernziel	Likert-Skala (1-5)	
Der gesamte Inhalt des Arbeitsblattes passt fachlich zum Lernziel.	5	5
Die Menge an Informationen im Arbeitsblatt passt zum Lernziel.	4	5
Der gesamte Inhalt des Arbeitsblattes ist fachlich angemessen für die Jahrgangsstufe.	3	4
Die Menge an Informationen im Arbeitsblatt ist angemessen für die Jahrgangsstufe.	4	4
Durchschnittsnote für Kategorie:	4.0	4.5
2. Inhalt des Arbeitsblattes	Likert-Skala (1-5)	
Im Text werden fachliche Konzepte und Theorien wiedergegeben.	4	4
Im Bild wird ein konkretes Beispiel wiedergegeben.	4	4
Das gesamte Arbeitsblatt ist inhaltlich konsistent und zusammenhängend.	3	4
Das Arbeitsblatt ist vollständig und enthält nur für das Lernziel relevante Informationen.	2	3
<u>Dekorative Bilder</u> werden vermieden.	2	5
Die Überschrift passt zum Inhalt des Arbeitsblattes.	4	4
<u>Querverweise</u> im Text auf die entsprechenden Bilder sind vorhanden	2	3
Durchschnittsnote für Kategorie:	2.6	3.4

3. Bildtypen im Arbeitsblatt (Welche Bildtypen sind in dem Arbeitsblatt zu finden?)		Ja/Nein	
Realistisches Bild		Ja	Ja
Schematisches Bild		-	-
Logisches Bild		-	-
Fachspezifisches Diagramm		-	-
4. Bildabfolgen aus mehreren Einzelbildern		Likert-Skala (1-5)	
Die Bezüge zwischen den Einzelbildern werden im Text, Abbildungstitel oder in einer Aufgabenstellung verdeutlicht.		4	5
Die Zusammenhänge zwischen den Bildern werden im Text, Abbildungstitel oder in einer Aufgabenstellung verdeutlicht.		5	5
Die einzelnen Bildelemente sind nummeriert oder ihre Reihenfolge ist anderweitig gekennzeichnet.		3	4
Durchschnittsnote für Thema:		4.0	4.7
5. Beschriftungen im Bild		Likert-Skala (1-5)	
Der Einsatz bzw. Nichteinsatz von Beschriftungen im Bild ist angemessen.		4	5
Die Beschriftungen im Bild sind kurz und prägnant.		3	4
Die Beschriftungen im Bild beziehen sich deutlich auf das relevante Element (z. B. durch Striche, Pfeile).		3	4
Durchschnittsnote für Thema:		3.3	4.3
6. Gestaltung der Abbildungen		Likert-Skala (1-5)	
Beim ersten Betrachten wird direkt deutlich, welche Bildelemente in der Abbildung am wichtigsten sind.		4	5
Zusammengehörige Bildelemente sind räumlich nahe.		4	5
Nicht zusammengehörige Bildelemente sind räumlich getrennt.		5	5
Die didaktisch intendierten Inhalte sind durch Hervorhebungen (z. B. Pfeil, unterschiedliche Farbe, Beschriftung) betont.		5	5
Durchschnittsnote für Thema:		4.5	5.0
7. Basistexte (Welche Arten von Basistexten sind im Arbeitsblatt zu finden?)		Ja/Nein	
Narrativer Text		-	-
Deskriptiver Text		Ja	Ja
Explikativer Text		-	Ja
Instruktiver Text		-	-
Argumentativer Text		-	-
8. Zusatztexte (Welche Arten von Zusatztexten sind im Arbeitsblatt zu finden?)		Ja/Nein	
Aufgaben		Ja	Ja
Lernzielangaben		-	-
Zusammenfassung		-	Ja
Exkurs		Ja	Ja
Glossar		-	-

9. Einleitungstext des Arbeitsblattes		Likert-Skala (1-5)	
Zu Beginn wird das Thema mit einem <u>Kontext</u> aufgegriffen.		3	4
Der Kontext wirft ein Problem, eine Fragestellung oder Aufgabe auf.		3	4
Der Kontext hat einen Bezug zur Alltags- oder Erfahrungswelt der Lernenden.		4	4
Der Kontext ist für das Lernziel und die Jahrgangsstufe angemessen.		2	3
Durchschnittsnote für Thema:		3.0	3.8
10. Gliederung und Ordnung des Textes		Likert-Skala (1-5)	
Der Lerngegenstand wird zuerst als Ganzes, dann detailliert betrachtet.		4	4
Es ist ein roter Faden klar erkennbar.		3	4
Der Gesamttext ist durch inhaltlich sinnvolle Absätze gegliedert.		3	4
Überschriften oder wichtige Begriffe sind <u>typographisch</u> hervorgehoben.		3	4
Unterschiedliche <u>Texttypen</u> sind z. B. durch Absätze, Umrahmungen klar voneinander abgegrenzt.		4	5
Aufzählungen sind nummeriert oder in Stichpunkten dargestellt.		1	3
Durchschnittsnote für Thema:		3.0	4.0
11. Sprachliche Angemessenheit des Textes		Likert-Skala (1-5)	
Der im Text beschriebene Inhalt ist fachlich korrekt.		3	3
Der Gesamttext ist fachsprachlich korrekt formuliert.		2	3
Der Textumfang ist für die Jahrgangsstufe angemessen.		2	4
Das sprachliche Niveau an Sätzen und Begriffen ist für die Jahrgangsstufe angemessen.		1	3
Die Satzstruktur ist für die Jahrgangsstufe angemessen.		1	3
Fachbegriffe werden einheitlich benutzt.		1	1
Durchschnittsnote für Thema:		1.7	2.8
12. Leserlichkeit des Textes		Likert-Skala (1-5)	
Die Schriftart, -größe und der Zeilenabstand sind leserlich und einfach gehalten.		1	3
Der Texthintergrund ist weiß.		1	3
Der Text enthält wenige Worttrennungen.		1	3
Die <u>Zeilenbreite</u> beschränkt sich auf 5-7 Wörter.		1	3
Durchschnittsnote für Thema:		1.0	3.0

13. Aufgabenstellungen des Arbeitsblatts		Likert-Skala (1-5)	
Die Aufgabenstellungen sind als konkrete Handlungsanweisungen formuliert.	1	3	
Die Aufgabenstellungen enthalten Operatoren.	1	5	
Es gibt Aufgabenstellungen zu unterschiedlichen Anforderungsbereichen.	1	4	
Die Aufgabenstellungen knüpfen in hohem Maß an die Informationen aus dem Text und/oder dem Bild an.	1	5	
In einer der Aufgabenstellungen werden die Schüler:innen damit beauftragt, Informationen aus Text und Bild zusammengetragen.	1	4	
In einer der Aufgabenstellungen werden die Schüler:innen damit beauftragt, die Inhalte aus Text und/oder Bild in eine neue Darstellungsform (z. B. Tabelle) umzuwandeln.	1	5	
Durchschnittsnote für Thema:	1.0	4.3	
14. In welcher Form liegt die Hauptrepräsentation vor?		Ja/Nein	
Der Text ist eine Hauptrepräsentation.	X	Ja	
Das Bild ist eine Hauptrepräsentation.	X	Ja	
15. In welcher Form liegt die Hilfsrepräsentation vor?		Ja/Nein	
Der Text ist eine Hilfsrepräsentation.	X	-	
Das Bild ist eine Hilfsrepräsentation.	X	-	
Es gibt keine Hilfsrepräsentation.	X	Ja	
16. Inhaltlicher Zusammenhang zwischen Text und Bild (Wie hängen Text und Bild inhaltlich zusammen?)		Ja/Nein	
Text und Bild helfen sich gegenseitig zum Verständnis.	Ja	Ja	
Es gibt eine einzige Hauptrepräsentation. Die Hilfsrepräsentation ist unterstützend und trägt zum Verständnis der Hauptrepräsentation bei.	-	-	
Durch die Hilfsrepräsentation können die Inhalte aus der Hauptrepräsentation vertieft und abstrahiert werden.	-	-	
17. Gegenseitige Unterstützung von Text und Bild		Likert-Skala (1-5)	
Text und Bild überschneiden sich inhaltlich.	2	2	
18. Ebenen von Repräsentationen (Welche Ebenen von Repräsentationen sind in dem Arbeitsblatt zu finden?)		Ja/Nein	
Makroskopische Ebene	Ja	Ja	
Mikroskopische Ebene	-	-	
Submikroskopische Ebene	Ja	Ja	
Symbolische Ebene	-	-	

19. Auswahl an Ebenen		Likert-Skala (1-5)	
Der Einsatz der ausgewählten Ebenen ist für das Lernziel angemessen.	1	1	
Die Ebenen sind in allen <u>Repräsentationstypen</u> klar voneinander getrennt.	1	4	
Die Ebenen im Text sind klar voneinander getrennt.	1	5	
Die Ebenen im Bild sind klar voneinander getrennt.	1	3	
Durchschnittsnote für Thema:	1.0	3.3	
20. Bewerten Sie folgende Aussage:		Ja/Nein	
Es findet mindestens ein Ebenenwechsel statt.	Ja	Ja	
21. Verknüpfung der Ebenen im Arbeitsblatt		Likert-Skala (1-5)	
Alle vorhandenen Ebenen nehmen Bezug auf die <u>makroskopische Ebene</u> .	1	5	
Die dargestellten Ebenen in der Abbildung sind klar voneinander abgegrenzt.	1	4	
In der Abbildung ist der Bezug zwischen den Ebenen hergestellt (z. B. Pfeil, Vergrößerung).	1	5	
Die dargestellten Ebenen im Text werden inhaltlich getrennt.	1	3	
Im Text ist der Bezug zwischen den Ebenen beschrieben.	1	3	
Die dargestellten Ebenen werfen keine falschen Schüler:innenvorstellungen auf.	1	4	
Die Darstellung von Teilchen und Teilchenhintergrund wirft keine falschen Schüler:innenvorstellungen auf.	1	5	
Es wird verdeutlicht, dass die eingesetzten Teilchenmodelle Modelle zur <u>Erklärung der submikroskopischen Realität</u> sind.	1	5	
Durchschnittsnote für Thema:	1.0	4.3	
22. Räumliche Anordnung von Text und Bild		Likert-Skala (1-5)	
Zugehörige Text- und Bildstellen sind räumlich nahe positioniert.	3	4	
Die Größe der Abbildung ist angemessen.	3	2	
Wenn die Repräsentationen nebeneinanderstehen, befindet sich die <u>Hauptrepräsentation</u> links.	1	1	
Wenn die Repräsentationen untereinanderstehen, befindet sich die <u>Hauptrepräsentation</u> oben.	1	1	
Durchschnittsnote für Thema:	2.0	2.0	
Teilnoten für			
Die Gestaltung der Abbildung(en):	3.9	4.7	
Die Gestaltung der Textteile:	1.9	3.6	
Die Vernetzung von Text und Abbildung:	1.3	3.2	
Gesamtnote für Ihr Arbeitsblatt:	2.3	3.7	

Rückmeldung zum Arbeitsblatt:	Note	
Herzlichen Glückwunsch, Ihr Arbeitsblatt sieht insgesamt sehr gut aus. Es ist sehr lernförderlich gestaltet. Optimierbare Aspekte sind kaum vorhanden.	1.0-1.9	
Glückwunsch, Ihr Arbeitsblatt ist insgesamt auf einem guten Niveau. Es ist lernförderlich und für Schüler:innen interesselweckend gestaltet.	2.0-2.9	
Sie sind auf dem richtigen Weg. Ihr Arbeitsblatt ist insgesamt auf einem befriedigenden Niveau gestaltet. Achten Sie etwas mehr auf fachliche Konventionen und fachdidaktische Grundprinzipien, die Sie bisher im Studium und im Fachseminar kennengelernt haben.	3.0-3.9	←
Ihr Arbeitsblatt ist insgesamt auf einem ausreichenden bis mangelhaften Niveau. Es beinhaltet leider viele Aspekte, die nochmal durchdacht und grundlegend überarbeitet werden müssten. Bitte achten Sie in der nächsten Sitzung vermehrt auf fachliche Konventionen und fachdidaktische Grundprinzipien, die Sie bisher im Studium und Fachseminar kennengelernt haben. In der nächsten Sitzung werden Sie sicherlich schon Fortschritte zeigen.	4.0-5.0	

Anmerkung. Skala/Antwortmöglichkeiten: 1 = trifft vollständig zu, 2 = trifft eher zu, 3 = trifft teilweise zu, 4 = trifft eher nicht zu, 5 = trifft gar nicht zu, X = weiß ich nicht; punktiert unterstrichene Begriffe sind im Glossar erläutert.

A.5 Test zur Erfassung der repräsentationsbezogenen Überzeugungen

Anmerkung für die statistische Auswertung. Items deren Itemnamen durchgestrichen sind, wurden in der Hauptstudie zwar eingesetzt, wurden aber während der statistischen Auswertungen aufgrund ihrer schlechten Kennwerte aussortiert. Darüber hinaus sind einige Items negativ formuliert und sollten vor den Berechnungen umgepolt werden. Dies gilt für die Items UBVE1, UBVE3, UBGE3, UBGE4.

Aufgabenbeschreibung:

Unten finden Sie allgemeine Aussagen über das Unterrichtsfach Chemie. Wie stehen Sie zu den Aussagen mit Ihrer persönlichen Meinung? Kreuzen Sie jeweils an (1 = Stimme gar nicht zu, 2 = Stimme eher nicht zu, 3 = Stimme teilweise zu, 4 = Stimme eher zu, 5 = Stimme voll zu). Da es sich um eine Meinungsabfrage handelt, gibt es keine richtigen oder falschen Antworten.

Itembeschreibungen:

1. Als Chemielehrkraft sollte ich ausreichende Kenntnisse haben, um einen bestimmten Sachverhalt (z. B. Säure-Base-Titration) durch unterschiedliche Darstellungsformen darbieten zu können (z. B. Text, Diagramm, Formel). [UBQU1]
2. Als Chemielehrkraft sollte ich Arbeitsblätter danach überprüfen, ob sie die für einen Unterrichtsgegenstand richtigen Darstellungsformen (z. B. Text, Diagramm, Tabelle, Formel) enthalten. [UBGE1]
3. Als Chemielehrkraft sollte ich im Unterricht möglichst allgemein bekannte Darstellungsformen einsetzen. [UBVE1]
4. Es ist Ziel des Chemieunterrichts, dass Schüler:innen mit fachtypischen Darstellungsformen kompetent umgehen. [UBZI1]
5. Als Chemielehrkraft sollte ich Darstellungsformen einsetzen, die an das Vorwissen der Schüler:innen anknüpfen. [UBVE2]
6. Als Chemielehrkraft sollte ich die Texte auf Arbeitsblättern explizit danach überprüfen, ob sie didaktisch gut formuliert und aufgebaut sind. [UBGE2]
7. Wenn Schüler:innen fachtypische Darstellungsformen gut beherrschen, dann lässt dies auf einen erfolgreichen Lernprozess schließen. [UBQU2]

8. Wenn es für einen Unterrichtsgegenstand eine allgemein bekannte Darstellungsweise gibt, so ziehe ich diese einer fachtypischen vor. [UBVE3]
9. Wenn ich einen Text aus einem Schulbuch für ein Arbeitsblatt nutzen möchte, brauche ich ihn in der Regel inhaltlich nicht zu überarbeiten. Ich kann ihn direkt übertragen. [UBGE3]
10. Es ist Ziel des Chemieunterrichts, dass Schüler:innen Skizzen, Diagramme und chemische Formeln selbstständig erstellen. [UBZI2]
11. Vorgefertigte Abbildungen, wie sie z. B. in Schulbüchern oder im Internet zu finden sind, eignen sich für eine Verwendung auf Arbeitsblättern. [UBGE4]
12. Verständlichkeit ist mir im Chemieunterricht wichtiger als eine exakte fachtypische Darstellungsweise. [UBVE4]
13. Es ist Ziel des Chemieunterrichts, dass Schüler:innen verschiedene Abbildungs-, Diagramm- Formeltypen lesen und beschreiben. [UBZI3]
14. Wenn ich für meinen Unterricht ein Liniendiagramm finde, die Skalierung der Achsen jedoch nicht passt, dann überarbeite ich das Liniendiagramm vor dem Einsatz. [UBGE5]
15. Als Chemielehrkraft sollte ich im Unterricht so häufig wie möglich unterschiedliche Darstellungsformen einsetzen. [UBVE5]
16. Es ist Ziel des Chemieunterrichts, Schüler:innen mithilfe von unterschiedlichen Darstellungsformen verschiedene Zugänge zu einem Unterrichtsinhalt darzubieten. [UBZI4]
17. Chemische Abbildungen sind für Schüler:innen lernförderlich, wenn Darstellungsebenen und Größendimensionen darin klar getrennt werden (z. B. Wassermoleküle herausgezoomt aus einem Becherglas). [UBGE6]
18. Als Chemielehrkraft sollte ich einen Unterrichtsgegenstand mit Hilfe von mehreren unterschiedlichen Darstellungsformen gleichzeitig darbieten (z. B. Text, Bild und Tabelle). [UBVE6]
19. Als Chemielehrkraft sollte ich meine Schüler:innen auch danach benoten, wie gut sie mit verschiedenen fachtypischen Darstellungsformen arbeiten können. [UBQU3]
20. Im Chemieunterricht sollten Schüler:innen lernen, Inhalte aus einer Darstellungsform (z. B. Tabelle) in eine andere Darstellungsform (z. B. Liniendiagramm) umzuwandeln. [UBZI5]

-
21. Als Chemielehrkraft sollte ich in meinem Unterricht Texte und Bilder einsetzen, die möglichst keine falschen Schüler:innenvorstellungen hervorrufen. [UBGE7]
22. Die Untersuchung von Aufbau und Funktion verschiedener Darstellungsformen sollte im Chemieunterricht ein eigener Unterrichtsgegenstand sein. [UBZI6]

B Erweiterte Ergebnisse aus der Pilotierung

Tabelle 51: Itemkennwerte zum Test des rep. Fachwissens in der Pilotierung

Item-name	Itemtrennschärfe (r_{it})	Reliabilität (α), wenn Item weggelassen	Popularitätsindex (p)	Kommentar
RVK01	.40	.72	.94	-
RTK01	.40	.7	.43	-
RIN05	.24	.70	.82	-
RKK03	.42	.72	.74	Kann in Moodle nicht umgesetzt werden (Aufgabe, in der gezeichnet werden soll)
RVK02	.61	.73	.70	Eher schlechte Kennwerte, aber inhaltlich wichtige Aufgabe
RTK02	.12	.71	.57	Aufgabe mit Zeichnen und sehr ähnlich zu RTK03; daher angepasst mit neuem Beispiel und ohne Zeichnen
RTN04	.15	.72	.67	Antworten aus halboffenem Aufgabenformat wurden genutzt, um eine Single-Choice-Aufgabe zu erstellen
RIK01	-.19	.70	.59	Schlechte Kennwerte
RKK02	.45	.70	.45	Kann in Moodle nicht umgesetzt werden (Aufgabe, in der gezeichnet werden soll)
RIN01	.37	.72	.37	Antworten aus halboffenem Aufgabenformat wurden genutzt, um eine Single-Choice-Aufgabe zu erstellen
RTK03	.33	.70	.28	-
RTN01	.13	.70	.34	Antworten aus halboffenem Aufgabenformat wurden genutzt, um eine Single-Choice-Aufgabe zu erstellen
RTN02	.28	.73	.30	Aufgabenformat angepasst
RIN03	.41	.73	.49	schlechte Kennwerte
RKK04	.40	.74	.61	schlechte Kennwerte
RKN01	.24	.70	.51	Kann in Moodle nicht umgesetzt werden (Aufgabe, in der gezeichnet werden soll)
RKK05	.42	.71	.14	-
RIN04	.61	.71	.83	bleibt wegen Inhaltsvalidität
RTN03	.12	.73	.81	Kann in Moodle nicht umgesetzt werden (Aufgabe, in der gezeichnet werden soll)

RVN01	.15	.72	.50	Kann in Moodle nicht umgesetzt werden (Aufgabe, in der gezeichnet werden soll)
------------------	-----	-----	-----	--

Anmerkung. $\alpha = .75$. Fettmarkierte Items wurden für die Hauptstudie beibehalten und durchgestrichene entfernt.

Tabelle 52: Itemkennwerte zum Test des rep. fachdidaktischen Wissen in der Pilotierung

Item-name	Itemtrennschärfe (r_{it})	Reliabilität (α), wenn Item weggelassen	Popularitätsindex (p)	Kommentar
REK01	.23	.57	.69	Geteilt in zwei Items
RVN02	.17	.58	.84	Bleibt wegen Inhaltsvalidität (Zusatzinformationen, dekorative Bilder)
REK03	.04	.60	.76	Keine inhaltliche Validität (Formeln), schlechte Itemtrennschärfe
REN13	.48	.52	.60	Gute Kennwerte, inhaltliche Validität (didaktische Hervorhebungen in Abbildungen)
REK02	.36	.56	.53	Aufgabenformat angepasst und Inhalt reduziert
RVN03	.09	.60	.19	Schlechte Itemtrennschärfe
RVN04	.14	.60	.50	schlechte Itemtrennschärfe, aber inhaltlich wichtige Aufgabe mit Teilchenmodell
REK04	.13	.59	.60	schlechte Itemtrennschärfe
RVN05	.10	.59	.44	schlechte Itemtrennschärfe
REN14	.24	.57	.44	-
REN05	.36	.55	.20	-
REN06	.35	.55	.43	-
REN07	.15	.58	.26	schlechte Itemtrennschärfe
REN12	.26	.57	.75	-
REN09	.13	.59	.53	-
RVN06	.32	.55	.56	-
REN11	.20	.58	.06	Kennwerte eher schlecht, aber hohe inhaltliche Validität (Identifizieren und Benennen von Ebenen)

Anmerkung. $\alpha = .63$. Fettmarkierte Items wurden für die Hauptstudie beibehalten und durchgestrichene entfernt.

Tabelle 53: Itemkennwerte zum Fragebogen der rep. Überzeugungen in der Pilotierung

Item- name	Itemtrenn- schärfe (r_{it})	Reliabilität (α), wenn Item weg- gelassen	Populari- tätsindex (p)
1	.12	.54	.69
2	.18	.53	.75
3	.20	.53	.87
4	.27	.51	.67
5	.40	.48	.59
6	.23	.52	.41
7	-.06	.58	.55
8	.08	.54	.89
9	.07	.55	.44
10	.08	.55	.70
11	.19	.53	.37
12	.45	.47	.65
13	.56	.46	.59
14	.23	.52	.66
15	.07	.55	.29
16	.01	.56	.73
17	.14	.54	.73

Anmerkung. $\alpha = .54$.

C Erweiterte Ergebnisse aus der Hauptstudie

Tabelle 54: Itemkennwerte zum Test des rep. Fachwissens im Posttest

Itemname	Itemtrennschärfe (r_{it})	Reliabilität (α), wenn Item weggelassen	Popularitäts- index (p)
1. RVK01	.44	.61	.66
2. RTK01	.18	.65	.47
3. RIN05	.78	.65	.82
4. RTN04	.31	.63	.63
5. RIN01a	.42	.61	.43
6. RIN01b	.34	.63	.39
7. RTK03	.19	.64	.28
8. RTN01	.12	.66	.62
9. RTN02	.27	.64	.87
10. RKK05	.28	.63	.22
11. RIN04	.29	.63	.81

Anmerkung. $\alpha = .65$.

Tabelle 55: Itemkennwerte zum Test des rep. fachdidaktischen Wissens im Posttest

Itemname	Itemtrennschärfe (r_{it})	Reliabilität (α), wenn Item weggelassen	Popularitäts- index (p)
1. REK05	.27	.62	.06
2. RVN02	.81	.63	.56
3. REN13	.24	.61	.36
4. REK02	.19	.62	.44
5. RVN04	.26	.61	.45
6. REN14	.14	.63	.36
7. REN05	.15	.63	.22
8. REN06	.23	.62	.47
9. REN12	.35	.60	.41
10. REN09	.34	.60	.27
11. RVN06	.35	.60	.56
12. REN11	.45	.59	.45

Anmerkung. $\alpha = .63$.

Tabelle 56: Itemkennwerte zum Fragebogen der rep. Überzeugungen im Posttest

	Itemname	Itemtrennschärfe (r_{it})	Reliabilität (α), wenn Item weggelassen	Popularitäts- index (p)
1.	UBZI1	.58	.76	.44
2.	UBZI2	.52	.79	.31
3.	UBZI3	.57	.79	.24
4.	UBZI4	.55	.79	.30
5.	UBZI5	.40	.80	.22
6.	UBZI6	.25	.81	.11
7.	UBVE1	.06	.82	.09
8.	UBVE2	.34	.80	.39
9.	UBVE3	-	-	-
10.	UBVE4	-	-	-
11.	UBVE5	.18	.81	.12
12.	UBVE6	.38	.80	.28
13.	UBGE1	.52	.79	.75
14.	UBGE2	.45	.80	.70
15.	UBGE3	-	-	-
16.	UBGE4	-	-	-
17.	UBGE5	.33	.80	.64
18.	UBGE6	.53	.79	.47
19.	UBGE7	.58	.79	.81
20.	UBQU1	.50	.80	.81
21.	UBQU2	.25	.81	.11
22.	UBQU3	.35	.80	.12

Anmerkung. $\alpha = .81$.

Tabelle 57: Ergebnisse der gruppenspezifischen Varianzanalysen der demografischen Daten

Variable	Shapiro-Wilk-Test [<i>p</i>]	Ergebnis der Analyse
Lehramt	-	$\chi^2(2) = 1.29, p = .524$
Schulform	-	$\chi^2(9) = 10.60, p = .390$
Form von angehender Lehrkraft	-	$\chi^2(2) = 3.37, p = .498$
Absolvierte Zeit	-	$\chi^2(6) = 5.49, p = .482$
Dauer als Vertretungslehrkraft	-	$\chi^2(6) = 2.53, p = .865$
Alter	-	$\chi^2(8) = 2.06, p = .979$
Abschlussnote Abitur	> .050	$F(2, 86) = 1.17, p = .314$
Abschlussnote Hochschule	< .050	$\chi^2(2) = .03, p = .984$
Muttersprache	-	$\chi^2(4) = 6.45, p = .168$

Tabelle 58: Ergebnisse der signifikanten dreifaktoriellen Varianzanalysen zu den Gesamtmittelwerten der Interventionsgruppe 1 aus allen drei Lehr-Lern-Materialien (Nr. 1, 2 und 3)

Kategorie	Ergebnis der Varianzanalyse	Bonferroni-korr. post-hoc Tests
Bild	$\chi^2(2) = 19.20, p < .001$	1/3: $z = 2.82, p = .014, \eta^2 = .27$ 2/3: $z = 3.87, p = .000, \eta^2 = .52$
Ebenen	$\chi^2(2) = 9.61, p = .008$	2/3: $z = 2.95, p = .009, \eta^2 = .30$

Tabelle 59: Ergebnisse der signifikanten dreifaktoriellen Varianzanalysen zu den Gesamtmittelwerten der Interventionsgruppe 2 aus allen drei Lehr-Lern-Materialien (Nr. 1, 2 und 3)

Kategorie	Ergebnis der Varianzanalyse	Bonferroni-korr. post-hoc Tests
Lernziel	$\chi^2(2) = 16.71, p < .001$	1/2: $z = -3.41, p = .002, \eta^2 = .40$ 1/3: $z = -3.48, p = .002, \eta^2 = .42$
Bild	$\chi^2(2) = 29.83, p < .001$	1/3: $z = 3.87, p = .000, \eta^2 = .52$ 2/3: $z = 4.60, p = .000, \eta^2 = .72$

Tabelle 60: Ergebnisse der signifikanten dreifaktoriellen Varianzanalysen zu den Gesamtmittelwerten der Kontrollgruppe aus allen drei Lehr-Lern-Materialien (Nr. 1, 2 und 3)

Kategorie	Ergebnis der Varianzanalyse	Bonferroni-korr. post-hoc Tests
Lernziel	$\chi^2(2) = 16.86, p < .001$	1/3: $z = -2.88, p = .012, \eta^2 = .27$ 1/2: $z = -3.69, p < .001, \eta^2 = .44$
Bild	$\chi^2(2) = 35.02, p < .001$	1/3: $z = 5.38, p < .000, \eta^2 = .94$ 2/3: $z = 4.00, p < .000, \eta^2 = .52$
Text	$\chi^2(2) = 11.44, p = .003$	1/3: $z = -2.81, p = .015, \eta^2 = .25$ 1/2: $z = -2.81, p = .015, \eta^2 = .25$
Einleitung	$\chi^2(2) = 37.28, p < .001$	1/3: $z = 2.44, p = .044, \eta^2 = .19$ 1/2: $z = 5.81, p < .000, \eta^2 = .88$ 2/3: $z = -3.38, p = .002, \eta^2 = .37$
Ebenen	$\chi^2(2) = 19.62, p < .001$	1/3: $z = 3.38, p = .002, \eta^2 = .37$ 2/3: $z = 4.13, p < .001, \eta^2 = .55$
Anordnung	$\chi^2(2) = 13.14, p = .001$	1/2: $z = -3.31, p = .003, \eta^2 = .35$

Tabelle 61: Prädiktoren des allgemeinen Fachwissens (prä) für die Gesamtstichprobe

Variable	<i>b</i>	<i>SE</i>	β	Korr. R^2	ΔR^2
Stufe 1				.14	.14
Rep. Fachwissen (prä)	.42	.11	.38***		
Stufe 2				.23	.09
Rep. Fachwissen (prä)	.36	.11	.33**		
Zeitdruck	10.33	3.30	.31**		
Stufe 3				.29	.06
Rep. Fachwissen (prä)	.33	.10	.30**		
Zeitdruck	9.31	3.22	.28**		
Selbstwirksamkeitserwartung	8.05	3.10	.25*		

Anmerkung. $F(82, 3) = 11.30, p < .001, N = 90$.

Tabelle 62: Prädiktoren des allgemeinen fachdidaktische Wissens (prä) in der Gesamtstichprobe

Variable	<i>b</i>	<i>SE</i>	β	Korr. R^2	ΔR^2
Stufe 1				.12	.12
Hochschulabschlussnote	8.59	2.45	.36***		
Stufe 2				.19	.07
Hochschulabschlussnote	8.28	2.35	.34**		
Allgemeines Fachwissen (prä)	.18	.06	.28**		

Anmerkung. $F(83, 2) = 10.77, p < .001, N = 90$.

Tabelle 63: Prädiktoren für das repräsentationsbezogene Fachwissen (prä) in der Gesamtstichprobe

Variable	<i>b</i>	<i>SE</i>	β	Korr. R^2	ΔR^2
Stufe 1				.13	.13
Allgemeines Fachwissen (prä)	.35	.09	.38***		
Stufe 2				.18	.05
Allgemeines Fachwissen (prä)	.28	.09	.31**		
Rep. Fachwissen (prä)	.29	.12	.24*		
Stufe 3				.21	.03
Allgemeines Fachwissen (prä)	.24	.09	.26*		
Rep. fachdidaktische Wissen (prä)	.29	.12	.24*		
Soziale Ressourcen	7.92	3.71	.21*		

Anmerkung. $F(82, 3) = 8.57, p < .001, N = 90$.

Tabelle 64: Prädiktoren für das repräsentationsbezogene Fachwissen (post) in der Gesamtstichprobe

Variable	<i>b</i>	<i>SE</i>	β	Korr. R^2	ΔR^2
Stufe 1				.34	.34
Rep. Fachwissen (prä)	.51	.08	.59***		
Stufe 2				.37	.04
Rep. Fachwissen (prä)	.49	.08	.56***		
Selbstwirksamkeitserwartung	5.51	2.25	.21*		
Stufe 3				.40	.03
Rep. Fachwissen (prä)	.54	.08	.62***		
Selbstwirksamkeitserwartung	5.73	2.21	.22*		
Rep. fachdidaktisches Wissen (prä)	.20	.09	.19*		

Anmerkung. $F(82, 3) = 19.73, p < .001, N = 90$.

Tabelle 65: Prädiktoren für das repräsentationsbezogene fachdidaktische Wissen (prä) in der Gesamtstichprobe

Variable	<i>b</i>	<i>SE</i>	β	Korr. R^2	ΔR^2
Stufe 1				.10	.10
Rep. Fachwissen (prä)	.28	.09	.34**		
Stufe 2				.14	.04
Rep. Fachwissen (prä)	.29	.09	.34***		
Elaborationsmuster	6.43	2.99	.22*		

Anmerkung. $F(83, 2) = 7.89, p < .001, N = 90$.

Tabelle 66: Prädiktoren für das repräsentationsbezogene fachdidaktische Wissen (post) aufgeteilt nach den Untersuchungsgruppen

Variable	<i>b</i>	<i>SE</i>	β	Korr. R^2	ΔR^2
Stufe 1, KG				.43	.43
Rep. fachdidaktisches Wissen (prä)	.63	.13	.67***		
Stufe 1, IG1				.21	.21
Rep. fachdidaktisches Wissen (prä)	.46	.16	.48*		
Stufe 1, IG2				.54	.54
Rep. fachdidaktisches Wissen (prä)	.86	.15	.75***		
Stufe 2, IG2				.64	.10
Rep. fachdidaktisches Wissen (prä)	.81	.13	.71***		
Selbstwirksamkeitserwartung	10.80	3.91	.32*		
Stufe 3, IG2				.71	.07
Rep. fachdidaktisches Wissen (prä)	.77	.12	.67***		
Selbstwirksamkeitserwartung	10.29	3.51	.31**		
Zeitdruck	7.66	2.88	.28*		

Anmerkung. KG: $F(27, 1) = 22.12, p < .001, n = 31$. IG1: $F(27, 1) = 8.24, p = .008, n = 29$. IG2: $F(24, 3) = 22.71, p < .001, n = 29$.

Tabelle 67: Prädiktoren für die repräsentationsbezogenen Überzeugungen (prä) in der Gesamtstichprobe

Variable	<i>b</i>	<i>SE</i>	β	Korr. R^2	ΔR^2
Stufe 1				.16	.16
Selbstwirksamkeitserwartung	.27	.07	.39***		
Stufe 2				.23	.07
Selbstwirksamkeitserwartung	.24	.07	.34***		
Hochschulabschlussnote	.18	.06	.27**		

Anmerkung. $F(83, 2) = 12.23, p < .001, N = 90$. Die Hochschulabschlussnoten sind umgepolt. Positive Regressionskoeffizienten sind daher als eine Verbesserung zu interpretieren.

Tabelle 68: Prädiktoren für die Qualität der designten Lehr-Lern-Materialien (Nr. 1) in der Gesamtstichprobe

Variable	<i>b</i>	<i>SE</i>	β	Korr. R^2	ΔR^2
Stufe 1				.10	.10
Allg. Fachwissen (prä)	.01	.00	.33**		
Stufe 2				.17	.07
Allg. Fachwissen (prä)	.01	.00	.32**		
Wichtigkeit von Wissenschaft	.23	.08	.28**		

Anmerkung. $F(83, 2) = 9.59, p < .001, N = 90$.

D Publikationen und weitere Beiträge

Veröffentlichungen

- Tonyali, B., Ropohl, M., & Schwanewedel, J. (2023). What makes representations good representations for science education? A teacher-oriented summary of significant findings and a practical guideline for the transfer into teaching. *Chemistry Teacher International*. Advance online publication. <https://doi.org/10.1515/cti-2022-0019>
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2023). *Einfluss von Feedback auf das Wissen von Lehramtsanwärter:innen*. In van Vorst, H. (Hrsg.), Lernen, Lehren und Forschen in einer digital geprägten Welt. Gesellschaft der Didaktik der Chemie und Physik, Jahrestagung in Aachen 2022 (Bd. 43, S. 194–197). Essen: Universität Duisburg-Essen. https://gdcp-ev.de/wp-content/uploads/securepdfs/2023/07/Tagungsband_2023_korrigiert.pdf
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2022). *Feedbackbasierte Fördermaßnahme zum Einsatz von Repräsentationen*. In Habig, S. & van Vorst, H. (Hrsg.), Unsicherheit als Element von naturwissenschaftsbezogenen Bildungsprozessen. Gesellschaft der Didaktik der Chemie und Physik, Virtuelle Jahrestagung 2021 (Bd. 42, S. 596–599). Essen: Universität Duisburg-Essen. <https://gdcp-ev.de/wp-content/uploads/2022/05/Tagungsband-2022-Stand-13522.pdf>
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2021). *Optimierung von Arbeitsblättern durch Feedback in Chemie*. In Habig, S. (Hrsg.), Naturwissenschaftlicher Unterricht und Lehrerbildung im Umbruch? Gesellschaft der Didaktik der Chemie und Physik online Jahrestagung 2020 (Bd. 41, S. 721–724). Essen: Universität Duisburg-Essen. https://www.gdcp-ev.de/wp-content/tb2021/TB2021_721_Tonyali.pdf
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2020). *Optimierung von Lehr-Lern-Materialien durch Feedback im Referendariat*. In S. Habig (Hrsg.), Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen. Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Wien 2019 (Bd. 40, S. 677–680). Essen: Universität Duisburg-Essen. https://www.gdcp-ev.de/wp-content/tb2020/TB2020_677_Tonyali.pdf

Tonyali, B., Ropohl, M. & Schwanewedel, J. (2019). *Feedback an Lehramtsanwärterinnen und -anwärter zum Einsatz externer Repräsentationen im naturwissenschaftlichen Unterricht*. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe*. Gesellschaft für Didaktik der Chemie und Physik (Bd. 39, S. 803–806). Jahrestagung in Kiel 2018. Regensburg: Universität Regensburg.
<https://gdcp-ev.de/?p=549>

Vorträge und Symposien

Tonyali, B., Ropohl, M. & Schwanewedel, J. (2025). *How to promote teachers' knowledge and self-assessment skills in designing instructional materials through internal feedback*. Oral presentation. ESERA Conference 2025, Kopenhagen.

Tonyali, B., Ropohl, M. & Schwanewedel, J. (2023). *Differential effects of internal and external feedback on different types of teachers' professional knowledge*. Oral presentation. NARST 2023 Annual International Conference, Vancouver.

Tonyali, B., Ropohl, M. & Schwanewedel, J. (2023). *Welches Feedback wirkt am besten? – Die Wirksamkeit von internem und externem Feedback bei Lehramtsanwärter:innen*. 10. GEBF-Jahrestagung 2023, Essen.

Tonyali, B., Ropohl, M. & Schwanewedel, J. (2022). *Welche Rolle spielen internes und externes Feedback bei der Förderung repräsentationsbezogenen Wissens von angehenden Lehrkräften?* Kolloquium. IZfB-Graduiertenkolloquium Naturwissenschaftlicher Unterricht, Essen.

Tonyali, B., Ropohl, M. & Schwanewedel, J. (2022) *Einfluss von Feedback auf das Wissen von Lehramtsanwärter:innen*. Vortrag. GDCP-Jahrestagung 2022, Aachen.

Tonyali, B., Ropohl, M. & Schwanewedel, J. (2022, akzeptiert). *Welchen Effekt haben internes und externes Feedback in einem Online-Lernmodul für angehende Lehrkräfte?* Vortrag. 9. GEBF-Jahrestagung 2022 (online), Bamberg.

Tonyali, B., Ropohl, M. & Schwanewedel, J. (2022). *Lehramtsanwärter*innen gestalten multiple externe Repräsentationen: Welches Feedback wirkt am besten?* Kolloquium. Neujahrskolloquium des Ortsverbands Duisburg-Essen der Gesellschaft Deutscher Chemiker, Essen.

-
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2021). *The effects of feedback on pre-service teachers' knowledge and beliefs about representations*. Oral presentation in a symposium. ESERA 2021 Online Conference.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2019). *Förderung des Einsatzes von Repräsentationen im Referendariat durch Feedback*. Kolloquium. Doktorierendenkolloquium der GDCP, Geiselwind.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2019). *Improving biology and chemistry pre-service teachers' use of external representations through feedback*. Oral presentation. ESERA Summer School, Chania.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2018). *Feedback an Lehramtsanwärterinnen und -anwärter zum Einsatz externer Repräsentationen im naturwissenschaftlichen Unterricht*. Kolloquium. IZfB-Graduiertenkolloquium Naturwissenschaftlicher Unterricht, Essen.

Posterbeiträge

- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2023). *Effekte von internem und externem Feedback auf das repräsentationsbezogene Wissen von Chemie-Lehramtsanwärter:innen*. Posterbeitrag. Tag der Bildungsforschung des Interdisziplinären Zentrums für Bildungsforschung (IZfB), Essen.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2021). *Feedbackbasierte Fördermaßnahme zum Einsatz von Repräsentationen*. Posterbeitrag. GDCP-Jahrestagung 2021 (Online).
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2020). *Optimierung von Arbeitsblättern durch Feedback in Chemie*. Posterbeitrag. GDCP-Jahrestagung 2020 (Online).
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2020, angenommen). *How To Give Effective Feedback To Pre-service Teachers About Their Representational Competences?* Round table poster paper. 2020 NARST Annual International Conference, Portland (cancelled), Oregon.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2020, angenommen). *Entwicklung und Evaluation eines Lernmoduls mit Feedback für angehende Lehrkräfte*. Posterbeitrag. 8. GEBF-Tagung 2020 (ausgefallen), Potsdam.

- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2020). *Entwicklung und Evaluation eines Lernmoduls mit Feedback für den Vorbereitungsdienst im Fach Chemie*. Posterbeitrag. Tag der Bildungsforschung des Interdisziplinären Zentrums für Bildungsforschung (IZfB), Essen.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2019). *Optimierung von Lehr-Lern-Materialien durch Feedback im Referendariat*. Posterbeitrag. GDGP-Jahrestagung, Wien.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2019). *Improving pre-service teachers' use of external representations through feedback*. Summer School poster. 2019 ESERA conference, Bologna.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2019). *Improving pre-service teacher's use of external representations through feedback*. Poster. ESERA Summer School, Chania.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2019). *Feedback an angehende Lehrkräfte zum Einsatz von externen Repräsentationen im naturwissenschaftlichen Unterricht*. Posterbeitrag. 7. GEBF-Jahrestagung 2019, Köln.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2019). *Feedback an angehende Lehrkräfte zum Einsatz von Repräsentationen*. Posterbeitrag. Tag der Bildungsforschung des Interdisziplinären Zentrums für Bildungsforschung (IZfB), Essen.
- Tonyali, B., Ropohl, M. & Schwanewedel, J. (2018). *Feedback an Lehramtsanwärterinnen und -anwärter zum Einsatz externer Repräsentationen im naturwissenschaftlichen Unterricht*. Posterbeitrag. GDGP-Jahrestagung 2018, Kiel.

Bisher erschienene Bände der Reihe
Studien zum Physik- und Chemielernen

ISSN 1614-8967

Vollständige Übersicht auf unserer Website



<https://www.logos-verlag.de/spcl>

Aktuelle Bände

- 311 Marvin Rost (2021): Modelle als Mittel der Erkenntnisgewinnung im Chemieunterricht der Sekundarstufe I. Entwicklung und quantitative Dimensionalitätsanalyse eines Testinstruments aus epistemologischer Perspektive
ISBN 978-3-8325-5256-5 44.00 EUR (open access)
- 312 Christina Kobl (2021): Förderung und Erfassung der Reflexionskompetenz im Fach Chemie
ISBN 978-3-8325-5259-6 41.00 EUR (open access)
- 313 Ann-Kathrin Beretz (2021): Diagnostische Prozesse von Studierenden des Lehramts. eine Videostudie in den Fächern Physik und Mathematik
ISBN 978-3-8325-5288-6 45.00 EUR (open access)
- 314 Judith Breuer (2021): Implementierung fachdidaktischer Innovationen durch das Angebot materialgestützter Unterrichtskonzeptionen. Fallanalysen zum Nutzungsverhalten von Lehrkräften am Beispiel des Münchener Lehrgangs zur Quantenmechanik
ISBN 978-3-8325-5293-0 50.50 EUR (open access)
- 315 Michaela Oettle (2021): Modellierung des Fachwissens von Lehrkräften in der Teilchenphysik. Eine Delphi-Studie
ISBN 978-3-8325-5305-0 57.50 EUR (open access)
- 316 Volker Brüggemann (2021): Entwicklung und Pilotierung eines adaptiven Multistage-Tests zur Kompetenzerfassung im Bereich naturwissenschaftlichen Denkens
ISBN 978-3-8325-5331-9 40.00 EUR (open access)
- 317 Stefan Müller (2021): Die Vorläufigkeit und soziokulturelle Eingebundenheit naturwissenschaftlicher Erkenntnisse. Kritische Reflexion, empirische Befunde und fachdidaktische Konsequenzen für die Chemielehrer*innenbildung
ISBN 978-3-8325-5343-2 63.00 EUR
- 318 Laurence Müller (2021): Alltagsentscheidungen für den Chemieunterricht erkennen und Entscheidungsprozesse explorativ begleiten
ISBN 978-3-8325-5379-1 59.00 EUR

- 319 Lars Ehlert (2021): Entwicklung und Evaluation einer Lehrkräftefortbildung zur Planung von selbstgesteuerten Experimenten
ISBN 978-3-8325-5393-7 41.50 EUR (open access)
- 320 Florian Seiler (2021): Entwicklung und Evaluation eines Seminarkonzepts zur Förderung der experimentellen Planungskompetenz von Lehramtsstudierenden im Fach Chemie
ISBN 978-3-8325-5397-5 47.50 EUR (open access)
- 321 Nadine Boele (2021): Entwicklung eines Messinstruments zur Erfassung der professionellen Unterrichtswahrnehmung von (angehenden) Chemielehrkräften hinsichtlich der Lernunterstützung
ISBN 978-3-8325-5402-6 46.50 EUR
- 322 Franziska Zimmermann (2022): Entwicklung und Evaluation digitalisierungsbezogener Kompetenzen von angehenden Chemielehrkräften
ISBN 978-3-8325-5410-1 49.50 EUR
- 323 Lars-Frederik Weiß (2021): Der Flipped Classroom in der Physik-Lehre. Empirische Untersuchungen in Schule und Hochschule
ISBN 978-3-8325-5418-7 51.00 EUR
- 324 Tilmann Steinmetz (2021): Kumulatives Lehren und Lernen im Lehramtsstudium Physik. Theorie und Evaluation eines Lehrkonzepts
ISBN 978-3-8325-5421-7 51.50 EUR
- 325 Kübra Nur Celik (2022): Entwicklung von chemischem Fachwissen in der Sekundarstufe I. Validierung einer Learning Progression für die Basiskonzepte „Struktur der Materie“, „Chemische Reaktion“ und „Energie“ im Kompetenzbereich „Umgang mit Fachwissen“
ISBN 978-3-8325-5431-6 55.00 EUR
- 326 Matthias Ungermann (2022): Förderung des Verständnisses von Nature of Science und der experimentellen Kompetenz im Schüler*innen-Labor Physik in Abgrenzung zum Regelunterricht
ISBN 978-3-8325-5442-2 55.50 EUR
- 327 Christoph Hoyer (2022): Multimedial unterstütztes Experimentieren im webbasierten Labor zur Messung, Visualisierung und Analyse des Feldes eines Permanentmagneten
ISBN 978-3-8325-5453-8 45.00 EUR
- 328 Tobias Schüttler (2022): Schülerlabore als interessefördernde authentische Lernorte für den naturwissenschaftlichen Unterricht nutzen
ISBN 978-3-8325-5454-5 50.50 EUR
- 329 Christopher Kurth (2022): Die Kompetenz von Studierenden, Schülerschwierigkeiten beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-5457-6 58.50 EUR
- 330 Dagmar Michna (2022): Inklusiver Anfangsunterricht Chemie. Entwicklung und Evaluation einer Unterrichtseinheit zur Einführung der chemischen Reaktion
ISBN 978-3-8325-5463-7 49.50 EUR

- 331 Marco Seiter (2022): Die Bedeutung der Elementarisierung für den Erfolg von Mechanikunterricht in der Sekundarstufe I
ISBN 978-3-8325-5471-2 66.00 EUR
- 332 Jörn Hägele (2022): Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten. Videobasierte Analysen zu Aktivitäten und Vorstellungen von Schülerinnen und Schülern der gymnasialen Oberstufe bei der Bearbeitung von fachmethodischer Instruktion
ISBN 978-3-8325-5476-7 56.50 EUR (open access)
- 333 Erik Heine (2022): Wissenschaftliche Kontroversen im Physikunterricht. Explorationsstudie zum Umgang von Physiklehrkräften und Physiklehramtsstudierenden mit einer wissenschaftlichen Kontroverse am Beispiel der Masse in der Speziellen Relativitätstheorie
ISBN 978-3-8325-5478-1 48.50 EUR (open access)
- 334 Simon Goertz (2022): Module und Lernzirkel der Plattform FLexKom zur Förderung experimenteller Kompetenzen in der Schulpraxis. Verlauf und Ergebnisse einer Design-Based Research Studie
ISBN 978-3-8325-5494-1 66.50 EUR
- 335 Christina Toschka (2022): Lernen mit Modellexperimenten. Empirische Untersuchung der Wahrnehmung und des Denkens in Analogien beim Umgang mit Modellexperimenten
ISBN 978-3-8325-5495-8 50.00 EUR (open access)
- 336 Alina Behrendt (2022): Chemiebezogene Kompetenzen in der Übergangsphase zwischen dem Sachunterricht der Primarstufe und dem Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-5498-9 40.50 EUR (open access)
- 337 Manuel Daiber (2022): Entwicklung eines Lehrkonzepts für eine elementare Quantenmechanik. Formuliert mit In-Out Symbolen
ISBN 978-3-8325-5507-8 48.50 EUR
- 338 Felix Pawlak (2022): Das Gemeinsame Experimentieren (an-)leiten. Eine qualitative Studie zum chemiespezifischen Classroom-Management
ISBN 978-3-8325-5508-5 46.50 EUR
- 339 Liza Dopatka (2022): Konzeption und Evaluation eines kontextstrukturierten Unterrichtskonzeptes für den Anfangs-Elektrizitätslehreunterricht
ISBN 978-3-8325-5514-6 69.50 EUR
- 340 Arne Bewersdorff (2022): Untersuchung der Effektivität zweier Fortbildungsformate zum Experimentieren mit dem Fokus auf das Unterrichtshandeln
ISBN 978-3-8325-5522-1 39.00 EUR (open access)

- 341 Thomas Christoph Münster (2022): Wie diagnostizieren Studierende des Lehramtes physikbezogene Lernprozesse von Schüler*innen?. Eine Videostudie zur Mechanik
ISBN 978-3-8325-5534-4 44.50 EUR (open access)
- 342 Ines Komor (2022): Förderung des symbolisch-mathematischen Modellverständnisses in der Physikalischen Chemie
ISBN 978-3-8325-5546-7 46.50 EUR
- 343 Verena Petermann (2022): Überzeugungen von Lehrkräften zum Lehren und Lernen von Fachinhalten und Fachmethoden und deren Beziehung zu unterrichtsnahem Handeln
ISBN 978-3-8325-5545-0 47.00 EUR (open access)
- 344 Jana Heinze (2022): Einfluss der sprachlichen Konzeption auf die Einschätzung der Qualität instruktionaler Unterrichtserklärungen im Fach Physik
ISBN 978-3-8325-5553-5 42.00 EUR (open access)
- 345 Jannis Weber (2022): Mathematische Modellbildung und Videoanalyse zum Lernen der Newtonschen Dynamik im Vergleich
ISBN 978-3-8325-5566-5 68.00 EUR (open access)
- 346 Fabian Sterzing (2022): Zur Lernwirksamkeit von Erklärvideos in der Physik. Eine Untersuchung in Abhängigkeit von ihrer fachdidaktischen Qualität und ihrem Einbettungsformat
ISBN 978-3-8325-5576-4 52.00 EUR (open access)
- 347 Lars Greitemann (2022): Wirkung des Tablet-Einsatzes im Chemieunterricht der Sekundarstufe I unter besonderer Berücksichtigung von Wissensvermittlung und Wissenssicherung
ISBN 978-3-8325-5580-1 50.00 EUR
- 348 Fabian Poensgen (2022): Diagnose experimenteller Kompetenzen in der laborpraktischen Chemielehrer*innenbildung
ISBN 978-3-8325-5587-0 48.00 EUR
- 349 William Lindlahr (2023): Virtual-Reality-Experimente. Entwicklung und Evaluation eines Konzepts für den forschend-entwickelnden Physikunterricht mit digitalen Medien
ISBN 978-3-8325-5595-5 49.00 EUR
- 350 Bert Schlüter (2023): Teilnahmemotivation und situationales Interesse von Kindern und Eltern im experimentellen Lernsetting KEMIE
ISBN 978-3-8325-5598-6 43.00 EUR
- 351 Katharina Nave (2023): Charakterisierung situativer mentaler Modellkomponenten in der Chemie und die Bildung von Hypothesen. Eine qualitative Studie zur Operationalisierung mentaler Modellkomponenten für den Fachbereich Chemie
ISBN 978-3-8325-5599-3 43.00 EUR
- 352 Anna B. Bauer (2023): Experimentelle Kompetenz Physikstudierender. Entwicklung und erste Erprobung eines performanzorientierten Kompetenzstrukturmodells unter Nutzung qualitativer Methoden
ISBN 978-3-8325-5625-9 47.00 EUR (open access)

- 353 Jan Schröder (2023): Entwicklung eines Performanztests zur Messung der Fähigkeit zur Unterrichtsplanung bei Lehramtsstudierenden im Fach Physik
ISBN 978-3-8325-5655-6 46.50 EUR
- 354 Susanne Gerlach (2023): Aspekte einer Fachdidaktik Körperpflege. Ein Beitrag zur Standardentwicklung
ISBN 978-3-8325-5659-4 45.00 EUR
- 355 Livia Murer (2023): Diagnose experimenteller Kompetenzen beim praktisch-naturwissenschaftlichen Arbeiten. Vergleich verschiedener Methoden und kognitive Validierung eines Testverfahrens
ISBN 978-3-8325-5657-0 41.50 EUR (open access)
- 356 Andrea Maria Schmid (2023): Authentische Kontexte für MINT-Lernumgebungen. Eine zweiteilige Interventionsstudie in den Fachdidaktiken Physik und Technik
ISBN 978-3-8325-5605-1 57.00 EUR (open access)
- 357 Julia Ortmann (2023): Bedeutung und Förderung von Kompetenzen zum naturwissenschaftlichen Denken und Arbeiten in universitären Praktika
ISBN 978-3-8325-5670-9 37.00 EUR (open access)
- 358 Axel-Thilo Prokop (2023): Entwicklung eines Lehr-Lern-Labors zum Thema Radioaktivität. Eine didaktische Rekonstruktion
ISBN 978-3-8325-5671-6 49.50 EUR
- 359 Timo Hackemann (2023): Textverständlichkeit sprachlich variiert physikbezogener Sachtexte
ISBN 978-3-8325-5675-4 41.50 EUR (open access)
- 360 Dennis Dietz (2023): Vernetztes Lernen im fächerdifferenzierten und integrierten naturwissenschaftlichen Unterricht aufgezeigt am Basiskonzept Energie. Eine Studie zur Analyse der Wirksamkeit der Konzeption und Implementation eines schulinternen Curriculums für das Unterrichtsfach „Integrierte Naturwissenschaften 7/8“
ISBN 978-3-8325-5676-1 49.50 EUR
- 361 Ann-Katrin Krebs (2023): Vielfalt im Physikunterricht. Zur Wirkung von Lehrkräftefortbildungen unter Diversitätsaspekten
ISBN 978-3-8325-5672-3 65.50 EUR (open access)
- 362 Simon Kaulhausen (2023): Strukturelle Ursachen für Klausurmisserfolg in Allgemeiner Chemie an der Universität
ISBN 978-3-8325-5699-0 37.50 EUR (open access)
- 363 Julia Eckoldt (2023): Den (Sach-)Unterricht öffnen. Selbstkompetenzen und motivationale Orientierungen von Lehrkräften bei der Implementation einer Innovation untersucht am Beispiel des Freien Explorierens und Experimentierens
ISBN 978-3-8325-5663-1 48.50 EUR (open access)
- 364 Albert Teichrow (2023): Physikalische Modellbildung mit dynamischen Modellen
ISBN 978-3-8325-5710-2 58.50 EUR (open access)

- 365 Sascha Neff (2023): Transfer digitaler Innovationen in die Schulpraxis. Eine explorative Untersuchung zur Förderung der Implementation
ISBN 978-3-8325-5687-7 59.00 EUR (open access)
- 366 Rahel Schmid (2023): Verständnis von Nature of Science-Aspekten und Umgang mit Fehlern von Schüler*innen der Sekundarstufe I. Am Beispiel von digital-basierten Lernprozessen im informellen Lernsetting Smartfeld
ISBN 978-3-8325-5722-5 53.50 EUR (open access)
- 367 Dennis Kirstein (2023): Individuelle Bedingungs- und Risikofaktoren für erfolgreiche Lernprozesse mit kooperativen Experimentieraufgaben im Chemieunterricht. Eine Untersuchung zum Zusammenhang von Lernvoraussetzungen, Lerntätigkeiten, Schwierigkeiten und Lernerfolg beim Experimentieren in Kleingruppen der Sekundarstufe I
ISBN 978-3-8325-5729-4 50.50 EUR (open access)
- 368 Frauke Düwel (2024): Argumentationslinien in Lehr-Lernkontexten. Potenziale englischer Fachtexte zur Chromatografie und deren hochschuldidaktische Einbindung
ISBN 978-3-8325-5731-7 63.00 EUR (open access)
- 369 Fabien Güth (2023): Interessenbasierte Differenzierung mithilfe systematisch variierten Kontextaufgaben im Fach Chemie
ISBN 978-3-8325-5737-9 48.00 EUR (open access)
- 370 Oliver Grewe (2023): Förderung der professionellen Unterrichtswahrnehmung und Selbstwirksamkeitsüberzeugungen hinsichtlich sprachsensibler Maßnahmen im naturwissenschaftlichen Sachunterricht. Konzeption und Evaluation einer video- und praxisbasierten Lehrveranstaltung im Masterstudium
ISBN 978-3-8325-5738-6 44.50 EUR (open access)
- 371 Anna Nowak (2023): Untersuchung der Qualität von Selbstreflexionstexten zum Physikunterricht. Entwicklung des Reflexionsmodells REIZ
ISBN 978-3-8325-5739-3 59.00 EUR (open access)
- 372 Dominique Angela Holland (2023): Bildung für nachhaltige Entwicklung (BNE) kooperativ gestalten. Vergleich monodisziplinärer und interdisziplinärer Kooperation von Lehramtsstudierenden bei der Planung, Durchführung und Reflexion von Online-BNE-Unterricht
ISBN 978-3-8325-5760-7 47.00 EUR (open access)
- 373 Renan Marcello Vairo Nunes (2024): MINT-Personal an Schulen. Eine Untersuchung der Arbeitssituation und professionellen Kompetenzen von MINT-Lehrkräften verschiedener Ausbildungswege
ISBN 978-3-8325-5778-2 51.00 EUR (open access)
- 374 Mats Kieserling (2024): Digitalisierung im Chemieunterricht. Entwicklung und Evaluation einer experimentellen digitalen Lernumgebung mit universeller Zugänglichkeit
ISBN 978-3-8325-5786-7 45.50 EUR
- 375 Cem Aydin Salim (2024): Die Untersuchung adaptiver Lernsettings im Themenbereich „Schwimmen und Sinken“ im naturwissenschaftlichen Unterricht
ISBN 978-3-8325-5787-4 49.00 EUR (open access)

- 376 Novid Ghassemi (2024): Evaluation eines Lehramtsmasterstudiengangs mit dem Profil Quereinstieg im Fach Physik. Erkenntnisse zu Eingangsbedingungen, professionellen Kompetenzen und Aspekten individueller Angebotsnutzung
ISBN 978-3-8325-5789-8 41.50 EUR (open access)
- 377 Martina Flurina Cavelti (2024): Entwicklung und Validierung eines Messinstruments zur Erfassung der Schülerkompetenzen im Bereich des wissenschaftlichen Skizzierens im Fach Chemie in der Sekundarstufe I
ISBN 978-3-8325-5829-1 45.00 EUR (open access)
- 378 Tom Bleckmann (2024): Formatives Assessment auf Basis von maschinellem Lernen. Eine Studie über automatisiertes Feedback zu Concept Maps aus dem Bereich Mechanik
ISBN 978-3-8325-5842-0 46.50 EUR (open access)
- 379 Jana Marlies Rehberg (2024): Das physikspezifische Mindset zum Studienbeginn. Fragebogenentwicklung und Aufbau einer Online-Intervention
ISBN 978-3-8325-5850-5 59.50 EUR (open access)
- 380 Florian Trauten (2024): Entwicklung und Evaluation von automatisierten Feedbackschleifen in Online-Aufgaben im Fach Chemie
ISBN 978-3-8325-5859-8 46.00 EUR (open access)
- 381 Johanna Dejanovikj (2024): Reflexion als Lerngegenstand im Unterricht zur Förderung von Bewertungskompetenz
ISBN 978-3-8325-5860-4 41.00 EUR (open access)
- 382 Katharina Flieser (2024): Verständlichkeit physikalischer Sachtexte. Untersuchungen zum Wirkungsgefüge zwischen sprachlicher Textgestaltung und der Behaltensleistung sowie der Textwahrnehmung im Schulfach Physik
ISBN 978-3-8325-5858-1 44.00 EUR (open access)
- 383 Stephanie Neppl (2024): Perspektivenübernahme im Physikunterricht. Explorative Interviewstudie zu einer Seminarkonzeption mit dem Schwerpunkt Perspektivenübernahme bei der Planung von Physikunterricht
ISBN 978-3-8325-5865-9 48.00 EUR (open access)
- 384 Katja Plicht (2024): Ein Physikübungskonzept zur Förderung der Problemlösekompetenz. Entwicklung und empirische Evaluation eines Strategietrainings auf der Basis von Expertisemerkmalen
ISBN 978-3-8325-5875-8 45.00 EUR (open access)
- 385 Svenja Boegel (2024): Feedback beim Experimentieren: Zur Rolle von Cognitive Load und Motivation
ISBN 978-3-8325-5911-3 45.50 EUR (open access)
- 386 Laura Pannullo (2025): Wahlmöglichkeiten beim Experimentieren: Entwicklung und Erprobung eines Konzeptes für Experimente in inklusiven Lerngruppen im Physikunterricht
ISBN 978-3-8325-5916-8 46.50 EUR (open access)

- 387 Matthias Schweinberger (2025): „Stumme Videos“. Aufmerksamkeitslenkende Moderation von Demonstrationsexperimenten
ISBN 978-3-8325-5930-4 50.50 EUR
- 388 Rike Große-Heilmann (2025): Entwicklung fachdidaktischen Wissens zum Einsatz digitaler Medien im Fach Physik
ISBN 978-3-8325-5943-4 51.00 EUR (open access)
- 389 Lion Cornelius Glatz (2025): Eine vergleichende Betrachtung der Überzeugungskraft von Experimenten zum Teilchenmodell
ISBN 978-3-8325-5954-0 42.00 EUR (open access)
- 390 Rasmus Viefers (2025): Förderung der Variablenkontrollstrategie im Sachunterricht
ISBN 978-3-8325-5968-7 44.50 EUR (open access)
- 391 Luc Broder Albrecht (2025): Kritisches Denken in der Chemielehrer*innenbildung. Eine konzeptionell-empirische Studie zur Integration und Operationalisierung
ISBN 978-3-8325-5970-0 57.00 EUR (open access)
- 392 Anna Weißbach (2025): Reflexion von Physikunterricht. Entwicklung und Validierung eines Testinstruments mit Assessment-Feedback zur Reflexionsfähigkeit Physik-Lehramtsstudierender
ISBN 978-3-8325-5977-9 51.50 EUR (open access)
- 393 Robert Gieske (2025): Förderung fachwissenschaftlicher und kommunikativer Kompetenzen von Schüler*innen durch Disaggregate Instruction im Fach Chemie. Konzeption und Durchführung einer vergleichenden Interventionsstudie zur Analyse der Wirksamkeit sprachsensibler Unterrichtsansätze in der Sekundarstufe I
ISBN 978-3-8325-5984-7 51.00 EUR (open access)
- 394 Thomas Sean Weatherby (2025): Talking Circuits. The Development and Assessment of a Digitally-Scaffolded, Collaborative Method for Teaching and Learning Electrical Circuits in Early Secondary Schools
ISBN 978-3-8325-6008-9 57.00 EUR (open access)
- 395 Julia Welberg (2025): Wen interessiert's? Quantitative Untersuchungen der Zusammenhänge der Neigungen zu empathisierender und systematisierender Denkweise und Interesse an Physik
ISBN 978-3-8325-6021-8 42.00 EUR (open access)
- 396 Sebastian Nickel (2026): Repräsentationskompetenzen im Fach Chemie. Eine empirische Untersuchung zur Dimensionalität des Konstrukts und zu dessen Zusammenhang mit Fachwissen und räumlichen Fähigkeiten
ISBN 978-3-8325-6074-4 55.00 EUR (open access)
- 397 Manuel Becker (2026): Der Energie-Feld-Ansatz: Entwicklung und Evaluation eines Konzeptes für den Energieunterricht der Sekundarstufe II
ISBN 978-3-8325-6079-9 53.50 EUR (open access)

- 398 Sevan Khagy (2026): Lernen durch Erklären – Entwicklung und Evaluation eines Seminarkonzepts zur Förderung der Sach- und Erklärkompetenz von Chemie-Lehramtsstudierenden
ISBN 978-3-8325-6081-2 52.50 EUR (open access)
- 399 Büşra Tonyalı (2026): Untersuchung der Bedingungen und Wirkungen von Feedback auf die professionelle Kompetenz von angehenden Lehrkräften im Vorbereitungsdienst zum Einsatz externer Repräsentationen
ISBN 978-3-8325-6112-3 46.50 EUR (open access)
- 400 Lotte Hahn (2026): Physik-Erklärvideos: Qualitätsanalyse und Erhebung von Einstellungen von (angehenden) Physiklehrkräften
ISBN 978-3-8325-6128-4 56.00 EUR (open access)

Vollständige Übersicht unter: <https://www.logos-verlag.de/spcl>

Alle erschienenen Bücher können unter der angegebenen ISBN direkt online (<http://www.logos-verlag.de>) oder telefonisch (030 - 42 85 10 90) beim Logos Verlag Berlin bestellt werden.

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf und Mathias Ropohl

Die Reihe umfasst inzwischen eine große Zahl von wissenschaftlichen Arbeiten aus vielen Arbeitsgruppen der Physik- und Chemiedidaktik und zeichnet damit ein gültiges Bild der empirischen physik- und chemiedidaktischen Forschung im deutschsprachigen Raum.

Die Herausgeber laden daher Interessenten zu neuen Beiträgen ein und bitten sie, sich im Bedarfsfall an den Logos-Verlag oder an ein Mitglied des Herausgeberteams zu wenden.

Kontaktadressen:

Univ.-Prof. Dr. Martin Hopf
Universität Wien,
Österreichisches Kompetenzzentrum
für Didaktik der Physik,
Porzellangasse 4, Stiege 2,
1090 Wien, Österreich,
Tel. +43-1-4277-60330,
e-mail: martin.hopf@univie.ac.at

Prof. Dr. Mathias Ropohl
Didaktik der Chemie,
Fakultät für Chemie,
Universität Duisburg-Essen,
Schützenbahn 70, 45127 Essen,
Tel. 0201-183 2704,
e-mail: mathias.ropohl@uni-due.de

Multiple externe Repräsentationen sind im Chemieunterricht ein wichtiges Medium der Wissensvermittlung. Befunde zeigen, dass Lehramtsstudierende über ein geringes repräsentationsbezogenes Wissen verfügen – ein Defizit, das auch im Vorbereitungsdienst fortbesteht. Allerdings benötigen Lehramtsanwärter:innen gerade in dieser Phase die entsprechende professionelle Kompetenz zur Unterrichtsplanung. Weitere Befunde lassen vermuten, dass Feedback ein wirksames Instrument zur Förderung der professionellen Kompetenz ist, insbesondere wenn internes und externes Feedback kombiniert werden.

Um die Wirksamkeit internen und externen Feedbacks auf die professionelle Kompetenz von Lehramtsanwärter:innen zu untersuchen, wurde eine Interventionsstudie mit $N = 90$ Proband:innen durchgeführt. Es wurde insbesondere der Einfluss auf das repräsentationsbezogene Fachwissen und repräsentationsbezogene fachdidaktische Wissen sowie auf die Überzeugungen analysiert.

Die Ergebnisse der Untersuchung zeigen, dass internes Feedback das repräsentationsbezogene fachdidaktische Wissen der Lehramtsanwärter:innen fördert. Die Kombination aus internem und externem Feedback führt nicht zu einem Wissenszuwachs, jedoch zu einer besseren Selbsteinschätzung. Die Überzeugungen haben sich als erwartbar stabil erwiesen.

Die Befunde bieten konkrete Hinweise zur Gestaltung und Optimierung von Ausbildungsformaten für Zentren für schulpraktische Lehrkräftebildung und Fachleitungen.

Logos Verlag Berlin

ISBN 978-3-8325-6112-3