

Kobl, Christina

Förderung und Erfassung der Reflexionskompetenz im Fach Chemie

Berlin : Logos Verlag 2021, III, 357 S. - (Studien zum Physik- und Chemielernen; 312) - (Dissertation, Universität Regensburg, 2020)



Quellenangabe/ Reference:

Kobl, Christina: Förderung und Erfassung der Reflexionskompetenz im Fach Chemie. Berlin : Logos Verlag 2021, III, 357 S. - (Studien zum Physik- und Chemielernen; 312) - (Dissertation, Universität Regensburg, 2020) - URN: urn:nbn:de:0111-pedocs-334164 - DOI: 10.25656/01:33416; 10.30819/5259

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-334164>

<https://doi.org/10.25656/01:33416>

in Kooperation mit / in cooperation with:



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

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden und es darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en> - You may copy, distribute and transmit, adapt or exhibit the work in the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work or its contents. You are not allowed to alter, transform, or change this work in any other way.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Studien zum Physik- und Chemielernen

M. Hopf, H. Niedderer, M. Ropohl, E. Sumfleth [Hrsg.]

312

Christina Kobl

Förderung und Erfassung der Reflexionskompetenz im Fach Chemie



λογος

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth

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, Hans Niedderer, Mathias Ropohl und Elke Sumfleth

Studien zum Physik- und Chemielernen

Band 312

Christina Kobl

Förderung und Erfassung der Reflexionskompetenz im Fach Chemie

Logos Verlag Berlin



Studien zum Physik- und Chemielernen

Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth [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.



© Copyright Logos Verlag Berlin GmbH 2021

Alle Rechte vorbehalten.

ISBN 978-3-8325-5259-6

ISSN 1614-8967

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>

Förderung und Erfassung der Reflexionskompetenz im Fach Chemie



Dissertation

ZUR ERLANGUNG DES AKADEMISCHEN GRADES
EINER DOKTORIN DER DIDAKTIK DER NATURWISSENSCHAFTEN
„DR. PHIL. NAT.“ (DOCTOR PHILOSOPHIAE NATURALIS)

Didaktik der Chemie

der Fakultät Chemie und Pharmazie
der Universität Regensburg

vorgelegt von
Christina Kobl (geb. Mathes)
aus Kelheim
2020

Vorsitzende: Prof. Dr. Astrid Rank
Erstgutachter: Prof. Dr. Oliver Tepner
Zweitgutachter: Prof. Dr. Stefan Krauss
Drittprüfer: Prof. Dr. Arne Dittmer
Tag der Disputation: 10.11.2020

Für die drei starken Männer in meinem Leben:

Papa, Oli und Hansi

Danke

Inhaltsverzeichnis

1.	Einleitung.....	1
2.	Unterrichtsqualität	3
2.1.	Modelle zum Konstrukt Unterrichtsqualität	3
2.2.	Unterrichtsqualitätsmerkmale	7
3.	Reflexionskompetenz	17
3.1.	Definitionen und Modelle zum Konstrukt Reflexion	17
3.1.1.	Definitionen.....	17
3.1.2.	Entwicklung der Modelle zum Konstrukt Reflexion	20
3.2.	Reflexion im Unterricht	30
3.2.1.	Reflexion und Profession	30
3.2.2.	Reflexionsfähigkeit – Voraussetzungen.....	51
3.3.	Förderung der Reflexionskompetenz	53
3.3.1.	Lehr-Lern-Labor.....	54
3.3.2.	Videografie.....	56
3.3.3.	Learning Study	59
3.3.4.	Feedback.....	62
3.4.	Messung der Reflexionskompetenz	68
3.4.1.	Modelle zur Messung der Reflexionskompetenz	68
3.4.2.	Wirksamkeit von Reflexionsprozessen	82
4.	Forschungsziele, Forschungsfragen und Hypothesen.....	91
4.1.	Ziele der Arbeit	91
4.2.	Forschungsfragen und Hypothesen.....	92
5.	Methodik und Durchführung	95
5.1.	Untersuchungsdesign und Stichprobengröße.....	95
5.2.	Durchführung der Studie.....	97
5.3.	Erhebungsinstrumente.....	103
5.3.1.	Testinstrumente und Fragebogen – Studierende	103
5.3.2.	Testinstrumente und Fragebogen – Schülerinnen und Schüler	113
5.4.	Datenauswertung.....	116
5.4.1.	Testgütekriterien.....	116
5.4.2.	Wissenstests.....	122
5.4.3.	Meinungsfragebogen	123
5.4.4.	Parametrische und nichtparametrische Tests	123
5.4.5.	Residuenberechnung	124
5.4.6.	Effektstärke	125

5.4.7. Korrelationen.....	125
5.4.8. Regressionsanalysen.....	126
5.4.9. Kodiermanual	126
6. Ergebnisse der Pilotstudie.....	129
6.1. Testinstrumente und Fragebogen – Studierende.....	129
6.1.1. Test zum fachdidaktischen Wissen	129
6.1.2. Fachwissenstest	131
6.1.3. Reflexionskompetenztest.....	133
6.1.4. Studierendenfragebogen.....	134
6.1.5. Kodiermanual	143
6.2. Testinstrumente und Fragebogen – Schülerinnen und Schüler.....	153
6.2.1. Wissenstest – Schülerinnen und Schüler.....	153
6.2.2. Kognitiver Fähigkeitstest – Schülerinnen und Schüler	156
6.2.3. Fragebogen zum Fachinteresse – Schülerinnen und Schüler	156
6.3. Konsequenzen für die Hauptstudie	157
7. Ergebnisse zur Hauptstudie	159
7.1. Ergebnisse zur Gestaltung eines universitären Seminars zur Förderung der Reflexionskompetenz (Forschungsfrage 1)	159
7.1.1. Ergebnisse zur Steigerung des deklarativen und konditionalen Wissens zum Thema Reflexionskompetenz (H 1)	159
7.1.2. Ergebnisse zur Steigerung des prozeduralen Wissens zum Thema Reflexionskompetenz (H1 und H 2).....	162
7.2. Ergebnisse der Zusammenhänge der Förderung der Reflexionskompetenz (Forschungsfrage 2)	170
7.2.1. Ergebnisse der Bereiche fachdidaktisches Wissen und Fachwissen sowie der Korrelationen bezüglich der Reflexionskompetenz mit den Bereichen fachdidaktisches Wissen und Fachwissen (H3a und H3b).....	170
7.2.2. Ergebnisse der Korrelationen bezüglich der Reflexionskompetenz und dem Studierendenfragebogen (H 4)	173
7.2.3. Ergebnisse des Wissenstests der Schülerinnen und Schüler über drei Messzeitpunkte (H 5)	179
7.2.4. Ergebnisse der Gruppenvergleiche im residualen Lernzuwachs (H6)	181
7.3. Weitere Ergebnisse	193
7.3.1. Experimentell-fachdidaktisches Wissen – Vergleich IG2 vs. KG	193
7.3.2. Korrelation Reflexionskompetenztest mit fachdidaktischem Wissen und Fachwissen unter Ausschluss der IG1 – Vergleich IG2 vs. KG	193
7.3.3. Kodiermanual – Perspektivenübernahme und Literaturbezug	193
7.3.4. Extremgruppenvergleiche	199

8.	Diskussion, Reflexion und Ausblick	201
8.1.	Gestaltung eines universitären Seminars zur Förderung der Reflexionskompetenz von Chemielehramtsstudierenden (Forschungsfrage 1)	201
8.2.	Zusammenhänge zwischen der Förderung der Reflexionskompetenz und diversen didaktischen Aspekten (Forschungsfrage 2)	202
9.	Literaturverzeichnis	213
10.	Anhang.....	244
10.1.	Verzeichnisse	244
10.1.1.	Abbildungsverzeichnis	244
10.1.2.	Tabellenverzeichnis.....	247
10.2.	Entwickelte Testinstrumente.....	261
10.2.1.	Reflexionskompetenztest.....	261
10.2.2.	Kodiermanual	266
10.3.	Statistiken.....	306
10.3.1.	Statistiken zum Kapitel 5.1	306
10.3.2.	Statistiken zum Kapitel 5.4.1	306
10.3.3.	Statistiken zum Kapitel 6.1.1	307
10.3.4.	Statistiken zum Kapitel 6.1.2	308
10.3.5.	Statistiken zum Kapitel 6.1.3	308
10.3.6.	Statistiken zum Kapitel 6.1.4	309
10.3.7.	Statistiken zum Kapitel 6.1.5	313
10.3.8.	Statistiken zum Kapitel 6.2.1	315
10.3.9.	Statistiken zum Kapitel 6.2.2	315
10.3.10.	Statistiken zum Kapitel 6.2.3.....	316
10.3.11.	Statistiken zum Kapitel 7.1.....	317
10.3.12.	Statistiken zum Kapitel 7.1.1.....	318
10.3.13.	Statistiken zum Kapitel 7.1.2.....	321
10.3.14.	Statistiken zum Kapitel 7.2.1.....	325
10.3.15.	Statistiken zum Kapitel 7.2.2.....	328
10.3.16.	Statistiken zum Kapitel 7.2.3.....	334
10.3.17.	Statistiken zum Kapitel 7.2.4.....	336
10.3.18.	Statistiken zum Kapitel 7.3.2.....	347
10.3.19.	Statistiken zum Kapitel 7.3.3.....	349
10.3.20.	Statistiken zum Kapitel 7.3.4.....	354

1. Einleitung

In den 70er Jahren kam der deutsche Bildungsrat zu dem Ergebnis, dass das Schlüsselproblem der Bildungsreform in der Lehrerbildung zu finden ist (Deutscher Bildungsrat, 1970). Es entwickelte sich in Deutschland eine Reformdebatte, zusätzlich angeheizt durch die Empfehlungen für die Lehrerbildung der Kultusministerkonferenz (KMK) 1998. Die aus Fachleuten aus Wissenschaft und Bildung zusammengesetzte Kommission unter Vorsitz von Ewald Terhart stellte, ausgehend von einer Situationsanalyse, Anforderungen für eine zukunftsorientierte Lehrerbildung vor (KMK, 2004). Weitere große Diskussionen lösten die Publikationen zu den nationalen und internationalen Schulvergleichsstudien wie PISA (Programme for International Student Assessment) (Baumert et al., 2004), TIMSS (Trends in International Mathematics and Science Study) (Kunter & Baumert) und IGLU (Internationale Grundschul-Lese-Untersuchung) (Bos, Tarelli, Bremerich-Vos & Schwippert, 2012) aus. Die detaillierten Ergebnisse dieser Studien in Bezug auf die Kompetenzen der Schülerinnen und Schüler in Deutschland ließen die Frage nach den entsprechenden Kompetenzen und der Professionalität der Lehrkräfte bzw. der Rolle der Lehrerinnen und Lehrer im Bildungssystem aufkommen. Rufe nach einer Neuausrichtung der Lehrerbildung wurden laut (Oelkers, 2007). So führte die Formulierung der Bildungsstandards für die Schule zu der Festlegung von Kompetenzbereichen mit zugeordneten Standards für die Lehrerbildung (KMK, 2004, S. 7–13).

Die Resultate der OECD-Lehrerstudie (Organisation for Economic Co-operation and Development) verwiesen auf ein mangelhaftes Bild in Bezug auf Struktur und Zielsetzung in der deutschen Lehrerbildung. Weitere Studien, wie zum Beispiel die COACTIV-Studie (Cognitive Activation in the Classroom – Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung mathematischer Kompetenz) zur empirischen Erfassung der professionellen Kompetenz von Lehrkräften oder die TEDS-M Studie (Teacher Education and Development Study: Learning to Teach Mathematics), eine internationale Vergleichsstudie zur Wirksamkeit der Lehrerbildung, folgten. Konsequenzen waren u. a. die Bologna-Reform, die zu einer Veränderung der Studienstruktur und einer Fokussierung auf die Kernpunkte führte. Als ein zentrales Element wurde die „fehlende Integration von Theorie und Praxis innerhalb der universitären Ausbildung“ gesehen (Hoppe-Graff, Schroeter & Flagmeyer, 2008, S. 354). Um den Übergang hierbei zu erleichtern und so auf die hohen gesellschaftlichen Anforderungen im Lehrberuf vorbereitet zu sein, bilden die Kompetenzen der KMK eine erste Orientierungshilfe. Ebenso werden Fördermaßnahmen aufgeführt, welche die Reflexion des eigenen Handelns und Planens umfassen (KMK, 2004, S. 6).

In den letzten Jahren hat die Förderung der Reflexionskompetenz in der Lehrerbildung an Bedeutung gewonnen. Eine Vielzahl an Studien belegen die positiven Auswirkungen der Reflexion im Zusammenhang mit verschiedenen Prozessen, wie unter anderem Selbstwahrnehmung, Metakognition und aktives Engagement, von denen wiederum erwartet wird, dass sie affektives und lebenslanges Lernen, aber auch Fähigkeiten wie Problemlösung, Kreativität und kritisches Denken fördern (Bain, Ballantyne, Packer & Mills, 1999; Boyd & Reese, 2006; Çimer, Çimer & Vekli, 2013; Estrada & Rahman, 2014; Hume, 2009; Loughran, 2002). Es wurden Konzepte entwickelt, die den Studierenden die Möglichkeit bieten, konkrete Situationen zu erleben und Erfahrungen zu machen, um so den schwierigen Transfer von Fähigkeiten und Theoriewissen auf den Beruf bewerkstelligen zu können (Bain et al., 1999; Çimer et al., 2013; Loughran, 2002). Lee (2008, S. 117) argumentiert: „It is only when [teacher candidates] reflect upon their knowledge critically that they can transfer what they have learned in initial teacher preparation programs as students to the real classroom situations as teachers.“

Um Forderungen nach einer reflexionsbasierten Ausbildung der Studierenden des Lehramts gerecht zu werden, wurde im Rahmen dieser Studie ein universitäres Seminar zur Förderung der Reflexionskompetenz entwickelt. Das Konzept sah vor, die Entwicklung des Wissens der Studierenden zum Thema Reflexion mithilfe eines selbst konzipierten Reflexionskompetenztests im Pre-Post-Follow-Up-Design abzubilden. Ebenso wurden Unterrichtsstunden entwickelt, gehalten und gleichzeitig videografiert, um diese im Anschluss mittels der Videos schriftlich zu reflektieren und die Stundenplanung dementsprechend überarbeiten zu können. Zur Auswertung der schriftlichen Reflexionen und zur Messung der Entwicklung der Reflexionskompetenz erfolgte der Einsatz eines selbst entwickelten Kodiermanuals. Dieser Ablauf wurde gemäß einer Learning Study (Nilsson, 2014) (Kapitel 3.3.3) und dem ALACT-Model (Korthagen & Vasalos, 2005) (Kapitel 3.1.2) einem Kreislauf folgend dreimal wiederholt. Um weitreichende Aussagen treffen zu können, wurden die Studierenden zusätzlich im Pre-Post-Follow-Up-Design auf ihr Fachwissen, ihr fachdidaktisches Wissen und ihre Selbstwirksamkeitserwartung hin untersucht. Ebenso ließen sich Rückschlüsse auf die Qualität der Unterrichtsstunde ziehen, indem das Fachwissen der Schülerinnen und Schüler vor und nach der Unterrichtsstunde getestet wurde.

In Kapitel 2 und 3 werden theoretische Grundlagen zur Unterrichtsqualität und zum Thema Reflexionskompetenz gelegt. Dabei wird Bezug auf aktuelle empirische Ergebnisse genommen. Kapitel 4 beinhaltet die Forschungsziele dieser Studie. Hieraus lassen sich Hypothesen ableiten, welche in Kapitel 4.2 dargestellt werden. Das darauffolgende Kapitel 5.1 beinhaltet das Untersuchungsdesign, die Stichprobengröße, die Erhebungsinstrumente und eine Erklärung des Vorgehens bei der Datenauswertung. Ergebnisse hierzu finden sich in Kapitel 6, beginnend mit der Pilotstudie und gefolgt von der Hauptstudie (Kapitel 7). Abschließend werden die Hypothesen in Kapitel 8 mit den Ergebnissen verglichen, diskutiert und unter Bezug zu ausgewählten methodischen Aspekten reflektiert.

2. Unterrichtsqualität

Das Kapitel Unterrichtsqualität ist unterteilt in zwei Unterkapitel. Teilkapitel 2.1. stellt nach einer kurzen Definition des Begriffs Unterrichtsqualität die Überlegungen von Lotz und Lipowsky (2015) sowie von Klieme und Rakoczy (2008) als allgemeine Modelle der Unterrichtsqualität vor. Es folgt eine spezifischere, auf den naturwissenschaftlichen Unterricht ausgelegte Ausführung von Fischer und Sumfleth (2013) und das häufig zitierte Modell von Helmke (2009). In Kapitel 2.2 wird zunächst der Begriff Unterrichtsqualitätsmerkmale definiert und im Zuge dessen werden eine Zusammensetzung allgemeiner Merkmale von Helmke (2009), ebenso wie unterschiedliche Auffassungen spezieller Merkmale des naturwissenschaftlichen Unterrichts vorgestellt.

2.1. Modelle zum Konstrukt Unterrichtsqualität

Der Begriff Unterrichtsqualität ist nach Einsiedler (2002) definiert als die Summe von Unterrichtsmerkmalen, die sich im Hinblick auf zuvor normativ festgesetzte Ziele als wirksam erwiesen haben. Diese Zielkriterien können psychosoziale und affektive Ziele sein, moralische Standards umfassen, aber auch reine Lernleistungen abbilden (Oser, 1992). Eine genauere Definition liefern Weinert, Schrader und Helmke (1989). Unterrichtsqualität ist demnach „jedes stabile Muster von Instruktionsverhalten, das als Ganzes oder durch einzelne Komponenten die substantielle Vorhersage und/oder Erklärung von Schulleistung erlaubt“ (Weinert et al., 1989, S. 899). In den letzten Jahren haben sich viele Modelle mit diversen Ansätzen aus unterschiedlichen Perspektiven zur Unterrichtsqualität entwickelt. Die Forschung unterscheidet das Persönlichkeitsparadigma, welches u. a. untersuchte, inwieweit die Lehrerpersönlichkeit Einfluss auf die Schülerleistung hat. Allerdings stellte sich dies als Sackgasse heraus (Bromme, 2014; Fischer & Sumfleth, 2013).

Weiter wurde das Prozess-Produkt-Paradigma untersucht, welches die Aspekte des Unterrichts bestimmte, die am höchsten mit verschiedenen empirisch gemessenen Einstellungs-, Interesse-, und Leistungsvariablen der Schülerinnen und Schüler korrelierten. Der letzte Forschungsansatz, das Expertenparadigma, identifizierte erfolgreiche Lehrkräfte und analysierte deren Unterricht. Das Themengebiet Professionswissen von Lehrkräften wurde stark von Shulman (1986, 1987) inspiriert.

Im deutschsprachigen Raum konnte das Modell der drei Basisdimensionen von Unterrichtsqualität (Abbildung 1) an Bedeutung gewinnen (Klieme, Lipowsky, Rakoczy & Ratzka, 2006).

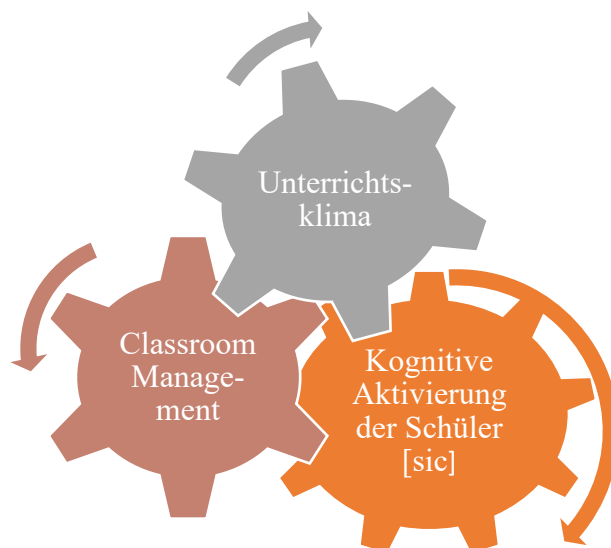


Abbildung 1: Drei Basisdimensionen der Unterrichtsqualität (Lotz & Lipowsky, 2015, S. 105)

Den großen Vorteil dieses Modells sehen Lotz und Lipowsky (2015) in der empirisch bewiesenen Grundlage der Basisdimensionen Unterrichtsklima, Classroom Management und Kognitive Aktivierung der Schülerinnen und Schüler. Zusätzlich sind die Einzelmerkmale guten Unterrichts (Kapitel 2.2) im Modell inkludiert, es repräsentiert somit einen großen Teilbereich der Unterrichtsgestaltung. Ebenso sehen die Autoren Vorteile in den faktorenanalytisch zusammengefassten Qualitätsmerkmalen, welche sich so weniger stark überlagern wie Merkmale anderer Konzeptualisierungen zur Unterrichtsqualität (Helmke, 2009; Meyer, 2003). Kritisiert werden hierbei die lediglich empirische Prüfung einzelner Merkmale und die fehlende Überprüfung deren Struktur. Klieme und Rakoczy (2008) ergänzen diese drei Basisdimensionen zu einem Modell (Abbildung 2), welches verdeutlicht, dass ein unterstützendes Unterrichtsklima primär motivationsfördernd ist, die kognitive Aktivierung zu einem systematischen Wissensaufbau und Verstehen führt und eine effektive Klassenführung, Regelklarheit und eine klare Struktur eine Voraussetzung für beide Wirkungspunkte bilden. Dies konnte in verschiedenen empirischen Studien bestätigt werden (Helmke & Klieme, 2008; Rakoczy et al., 2007):

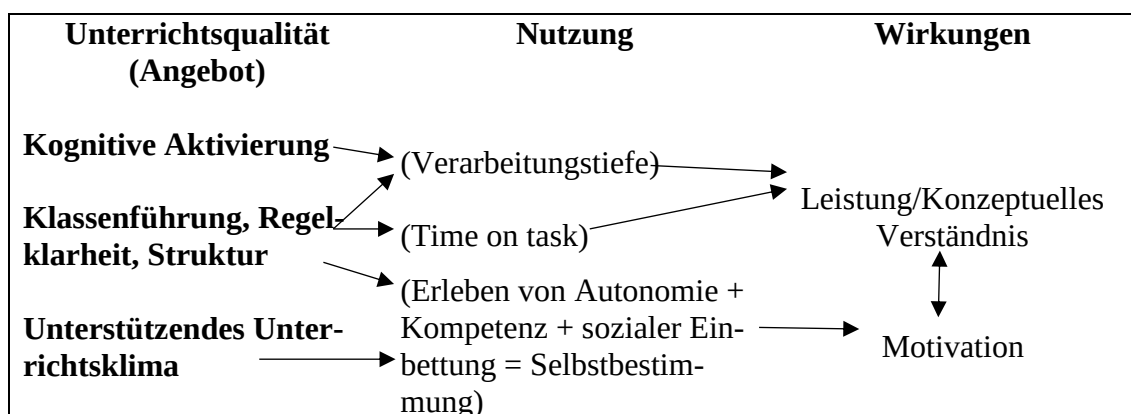


Abbildung 2: Grunddimensionen der Unterrichtsqualität und deren vermutliche Wirkung (Klieme & Rakoczy, 2008, S. 228)

Für den naturwissenschaftlichen Unterricht zeigt Abbildung 3 einen Teil möglicher Variablen, welche sich auf die drei Säulen Lehrperson, Lernende sowie Unterrichtsebene

und zusätzlich den möglichen Output stützt. Aus verschiedensten Studien ging hervor, dass das Zusammenspiel von Lernenden und Lehrpersonen ein wichtiges Element für die Beurteilung der Unterrichtssituation darstellt. Zusätzlich müssen aber Variablen wie Vorwissen, Kompetenzen, Motivation/Interesse, Lernstrategien und Lernerfolg auf Schülerinnen- und Schülerseite, sowie Aspekte wie Professionswissen, Lehrerenthusiasmus und Arbeitsbedingungen seitens der Lehrkräfte mit erhoben werden, um empirisch fundierte Aussagen über die Unterrichtsqualität machen zu können (Fischer & Sumfleth, 2013).

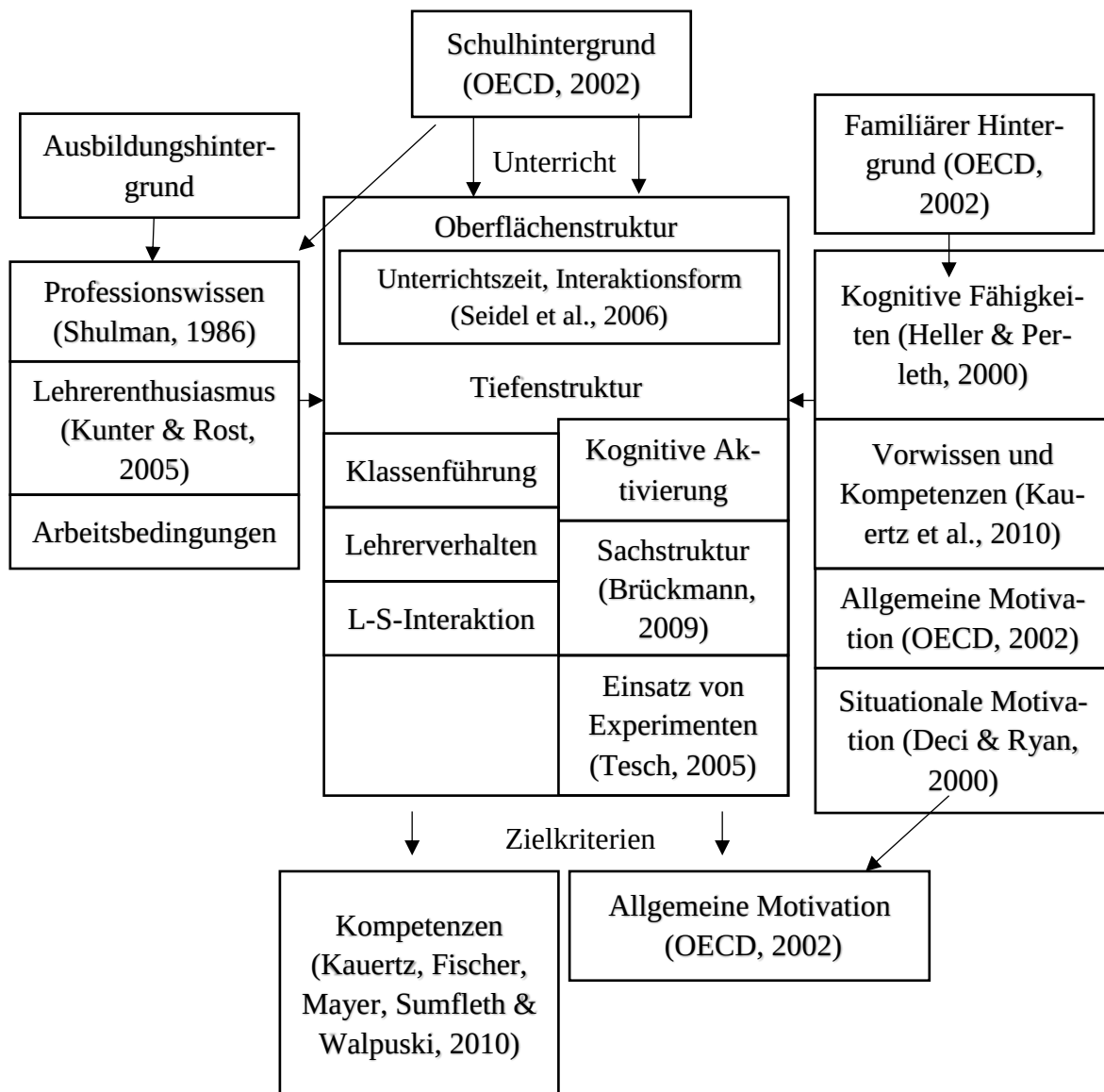


Abbildung 3: Modell für Unterrichtsqualität in den Naturwissenschaften mit exemplarischen Messskalen (H. Fischer & Sumfleth, 2013, IX)

Eines der meist zitierten und verwendeten Modelle ist das Angebots-Nutzungs-Modell (Abbildung 4) der Unterrichtswirksamkeit von Helmke (2009). Er erarbeitete dieses Modell auf Grundlage der Überlegungen von Weinert, welcher schrieb, dass es „isolierte, einfache, stabile und invariant gültige Abhängigkeitsbeziehungen zwischen Kriterien des Unterrichtserfolgs und Merkmalen des Unterrichts nicht gibt. Damit fehlt den vielen vereinfachenden Prozess-Produkt-Modellen des Lehrens die wissenschaftliche Basis“ (Weinert, 1999, S. 211). Diese Gedanken und weitere Einwände von Weinert und die theoretischen Überlegungen von Fend (1981) bildeten die Basis des Angebots-Nutzungs-

Modells von Helmke (2007b, 2009), welches versucht, „Faktoren der Unterrichtsqualität in ein umfassenderes Modell der Wirkungsweise und Zielkriterien des Unterrichts zu integrieren“ (Helmke, 2009, S. 73). Das auch von Helmke immer wieder überarbeitete und angepasste Modell implementiert sowohl Merkmale der Lehrperson als auch des Unterrichts:

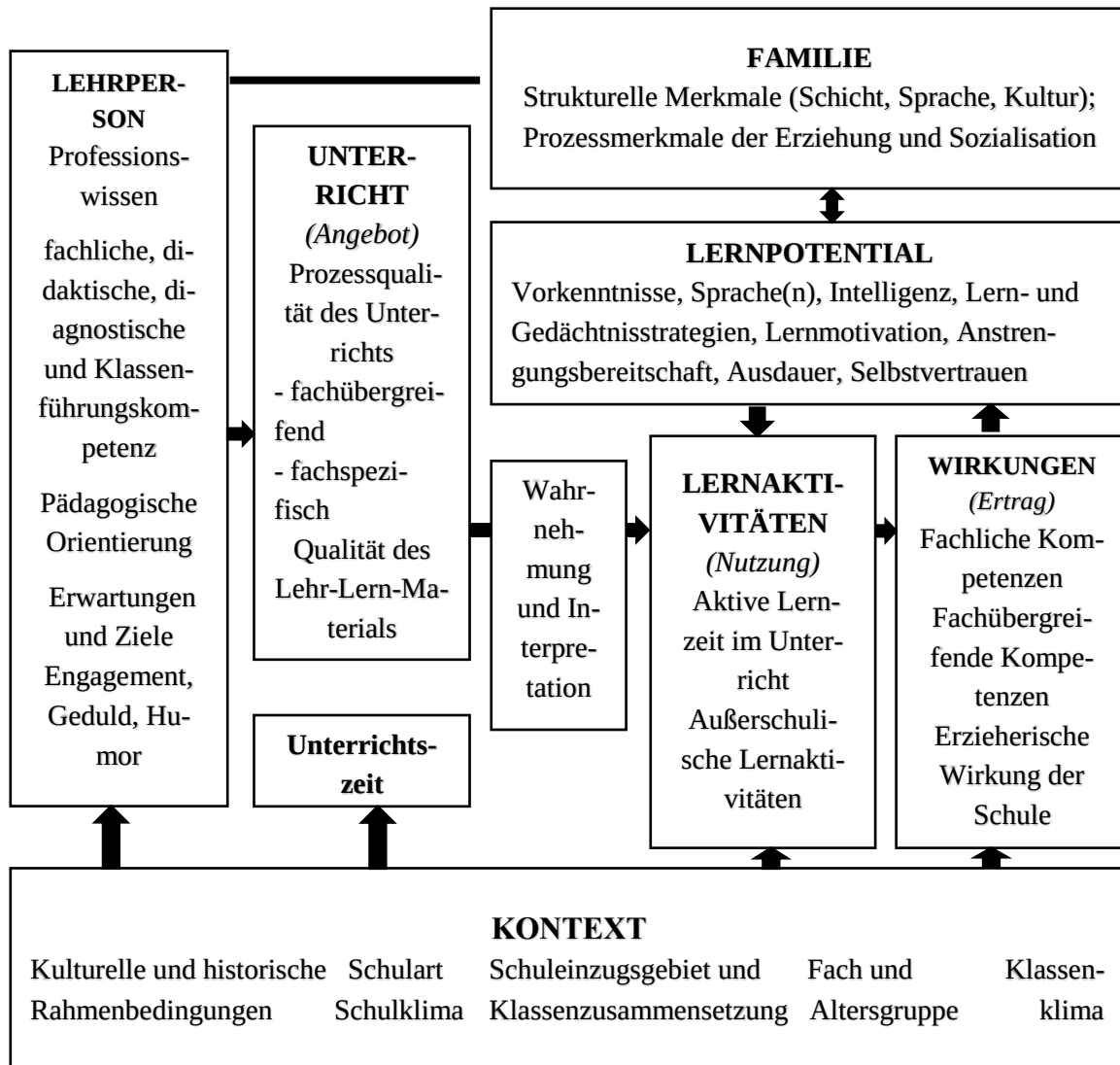


Abbildung 4: Angebots-Nutzungs-Modell (Helmke, 2012, S. 71)

Basis hierbei sind drei Säulen: Zunächst Unterricht als Angebot, welches von den Schülerinnen und Schülern in irgendeiner Weise genutzt werden muss, um bestimmte Ergebnisse (fachliche oder fachübergreifende Kompetenzen) erreichen zu können. Der Einsatz unterrichtlicher oder außerunterrichtlicher Lernaktivitäten entspricht der zweiten Säule, der Nutzung. Nur wenn diese stattfindet, kann eine angestrebte Wirkung, also ein Ertrag und somit die dritte Säule, resultieren. Ebenso haben Kontextmerkmale und individuelle Merkmale der Schülerinnen und Schüler Einfluss auf die Lernaktivität. Folglich sind Personenmerkmale der Lehrkraft in diesem Modell nur indirekt wirksam. Die entscheidende Rolle im unterrichtlichen Lernprozess spielen die Schülerinnen und Schüler selbst. Somit folgt, dass sich Unterrichtsangebote nie eins-zu-eins in Lernerfolge umwandeln lassen. Zentral bedeutsam im Modell ist das Wahrnehmungs- und Interpretationsfeld, welches Gelegenheit für Reflexionen auf Lehrer- und Schülerseite bietet (Kohler & Wacker,

2013). Dieses von Helmke selbst weiterentwickelte Modell operiert weitgehend mit Begriffen aus der Lehr-Lernforschung (Gruschka, 2011) und konzeptualisiert nach Eickhorst (2011) wichtige Gestaltungsmerkmale von Unterricht. Zusätzlich werden diese durch messbare Kriterien beschrieben (Terhart, 2009). Das Modell wurde auch von zahlreichen Forschern aufgegriffen und adaptiert (Lipowsky, 2015, S. 80 ff.; Reusser, Pauli & Waldis, 2010, S. 18). Lipowsky (2015) stellt in seiner vereinfachten Version des Angebots-Nutzungs-Modells den Klassenkontext als zentrales Element heraus und kann dies auch empirisch begründen. Das im Kern von Helmke übernommene Modell von Reusser et al. (2010) ergänzt zusätzlich die Inhaltsqualität des Lehrangebots. Hier wird Unterricht als Gesamtgefüge dargestellt, dessen Elemente zusammenhängen und nicht trennbar sind. Allgemein reduzieren Modelle wie diese die Komplexität von Vorgängen, um Analysen, Reflexionen wie auch Verständnis zu erleichtern. Das Angebots-Nutzungs-Modell kann jedoch als Basis und Orientierungshilfe für guten Unterricht dienen. Aus diesem Grund werden einzelne Punkte des Modells, ergänzt durch ausgewählte andere Modelle, im Folgenden näher erläutert.

2.2. Unterrichtsqualitätsmerkmale

Unter dem Begriff Unterrichtsqualitätsmerkmale verstehen Helmke und Schrader (2010, S. 27) „Aspekte, mittels derer sich Unterricht und Lehrer-Schüler-Interaktion beschreiben lassen. Diese Merkmale sind nicht direkt beobachtbar, sondern Abstraktionen. [...] Wenn hier von Merkmalen gesprochen wird, sind Variablen gemeint, also Größen, die in verschiedenen Ausprägungen variieren.“

In der Unterrichtsqualitätsforschung haben sich über die Jahre variierende Listen an Unterrichtsqualitätsmerkmalen herauskristallisiert. Welche Merkmale hierbei am wichtigsten sind, hängt dabei vom betrachteten Bildungsziel ab. Die empirische Bildungsforschung konnte aber nachweisen, dass einige Schlüsselmerkmale der Unterrichtsqualität existieren. Dies sind Klassenführung und Zeitznutzung, Schülerorientierung, lernförderliches Klima und Motivierung, Klarheit, Verständlichkeit und Strukturierung, Übung, Wiederholung und Konsolidierung, Aktivierung, Umgang mit Heterogenität (der Lernvoraussetzungen) und angemessene Methodenvariation sowie Wirkungs- und Kompetenzorientierung (Helmke, Goellner, Kleinbub, Schrader & Wagner, 2008). Ähnliche Faktoren für den Unterricht wie die bereits genannten kann auch Hattie (2008) in seiner Studie, welche zahlreiche Metaanalysen zur Unterrichtsqualität und zu erfolgreichem Lernen zusammenfasst, berichten. Zusätzlich zu nennen bleiben Zielorientierung, Lösungsbeispiele, Feedback, metakognitive Strategien und Problemlösen. Ebenso fand er, bezogen auf das Elternhaus, Faktoren wie den sozioökonomischen Index und die häusliche Umgebung als lernerfolgsfördernd. Auf Schülerinnen- und Schülerseite werden als Prädiktoren für Lernerfolg die Piaget'sche Entwicklungsstufe, das Vorwissen und die vorherigen Noten genannt. Auf Lehrkräfteseite haben Microteaching-Situationen in der Ausbildung, das Lehrer-Schüler-Verhältnis und die professionelle Entwicklung einen hohen Stellenwert.

Betrachtet man die in der Literatur aufgeführten Merkmale guten Unterrichts (Brophy, Good & Wittrock, 1986; Helmke, 2009; Tobin & Fraser, 2003), so findet man hauptsächlich obig erstgenannte Merkmale des Unterrichtsprozesses (Schulz, 2011). Diese Auflistung von Merkmalen korreliert grundsätzlich positiv mit dem Lernerfolg der Schülerinnen und Schüler oder der wahrgenommenen Unterrichtsqualität (Walpuski, 2017). Allerdings sind diese fachübergreifenden Kriterien nur teilweise über verschiedene Fächer

stabil. Walpuski (2017) weist darauf hin, dass bestimmte Qualitätsmerkmale eine fachabhängige Wirkung entfalten. So konnte z. B. gezeigt werden, dass sich die inhaltliche Strukturierung im mathematisch-naturwissenschaftlichen Unterricht als wirksamer erweist als beispielsweise im Englischunterricht (Klieme et al., 2006). Einen zusätzlichen Einfluss sieht Walpuski (2017) in der Art der Umsetzung durch die Lehrkraft und die Schülerinnen und Schüler in einer bestimmten Lernsituation. Betrachtet man die fachspezifische Sichtweise bzw. die spezifischen Eigenarten und Prozesse des Faches Chemie, dann stellt gerade das Experiment eine solche Herausforderung dar (Bader & Schmidkunz, 2002; Hofstein & Lunetta, 1982; KMK, 2004). Von Interesse ist hierbei die Studie von Schulz (2011), deren Ziel es war, experimentierspezifische Qualitätsmerkmale im Chemieunterricht in Bezug auf Leistung, Interesse und Motivation der Schülerinnen und Schüler zu identifizieren. Die Qualitätsmerkmale, welche am höchsten mit dem Lernerfolg korrelierten, wurden als mögliche positive Merkmale von Experimentierphasen in einer zweiten Interventionsstudie überprüft. Dabei konnten insbesondere Merkmale wie Schülererklärungen, hypothesenüberprüfendes Experimentieren in Kleingruppen, Instruktionseffizienz und problemlösender Unterricht als besonders lernwirksam gezeigt werden. Ebenso ist ein sehr guter Überblick über die bisherigen Ergebnisse der Qualitätsmerkmale naturwissenschaftlichen Unterrichts inkludiert (Schulz, 2011). Tabelle 1 gibt einen Überblick über ausgewählte Arbeiten zu Qualitätsmerkmalen naturwissenschaftlichen Unterrichts, welche auch in dieser Studie von Bedeutung sind:

Tabelle 1: Qualitätsmerkmale im naturwissenschaftlichen Unterricht (Schulz, 2011, S.13-14)

Autor / Quelle	Qualitätsmerkmale naturwissenschaftlichen Unterrichts
Fraser (1998)	- Gutes Klassenklima - Zielführender Lehrstil - Gute Unterrichtsorganisation - Kooperatives Lernen - Individualisierte Lernumgebungen - Unterstützung durch die Lehrkraft - Aufgabenorientierung - Anweisungen und Organisation
Harlen (1999)	- Eigene Erfahrung der Schüler durch praktisches Arbeiten unterstützen - Idee sollte anhand von Vorhersagen der Ergebnisse getestet werden - Experiment sollte auf Ideen abgestimmt sein - Experiment sollte theoriebasiert sein - Praktische Erfahrung durch Erkundung in den Naturwissenschaften sollte einen bestimmten Grad an Offenheit haben, besonders am Ende
Berg, Bergendahl, Lundberg und Tibell (2003) <i>[Anm. der Autorin: Die Autorennamen wurden gemäß</i>	- Offenes Experimentieren (mit Anleitung zum Vorgehen): - Selbstständiges Formulieren von Hypothesen - Selbstständige Planung des Experiments zur Überprüfung der Hypothesen - Selbstständige Durchführung des Experiments - Selbstständige Auswertung und Diskussion des Experiments

<i>der Originalstudie korrigiert.]</i>	
Clausen, Reusser und Klieme (2003)	- Instruktionseffizienz - Gute Klassenführung - Störungsfreie Unterrichtsabläufe - Hohes Ausmaß an Regelklarheit - Hohes Ausmaß an genutzter Unterrichtszeit - Schülerorientierung - Individualisierende Unterrichtsformen - Positive Schülerorientierung - Umgang mit Fehlern
Rumann (2005)	- Kooperatives Experimentieren in Kleingruppen
Walpuski (2006)	- Offenheit der Idee, Planung, Durchführung und Auswertung von Experimenten - Feedback und Fehlerkorrektur durch responsive Lehrerintervention
Wahser (2008)	- Strukturierungshilfen/Strukturierungstraining im Hinblick auf das strukturierte experimentelle Arbeiten
Wirth, Thillmann, Künsting, Fischer und Leutner (2008)	- Schülerexperimente mit vorgegebenen Lernzielen statt Problemlösezielen und unterstützende Regulation des Lernprozesses durch metakognitive Prompts
Wüsten (2010)	- Effektive Klassenführung - Klarheit der Organisation - Schüleraktivität - Schülerfragen

Im Folgenden werden die in dieser Studie besonders relevanten Qualitätsmerkmale näher erläutert.

Klassenführung (Classroom Management)

Klassenführung lässt sich definieren als alles,

„was Lehrpersonen mittels Aktivitäten und Haltungen zur Steuerung der Interaktionen in der Klasse beitragen, wobei ihnen bewusst ist, dass die Klasse mehr ist als die Summe der einzelnen Schüler und dass sich die individuellen und die sozialen Lernprozesse gegenseitig beeinflussen. Ziel ist ein Klassenklima, welches gute Lehr- und Lernprozesse ermöglicht, die Entfaltung und den Schutz jedes Einzelnen gewährleistet, den Schülern ermöglicht, aus gemeinschaftsbildenden Aktivitäten zu lernen und die Motivation der Schüler stärkt, sich zugunsten der Gemeinschaft einzusetzen“ (Bildungsdirektion Kanton Zürich, 2006, S. 1).

Weinert und Helmke (1996) sehen als zentrale Elemente der Klassenführung, die Schülerinnen und Schüler zu motivieren, den Unterricht möglichst störungsarm zu gestalten und auftretende Störungen schnell zu beenden. Besonders wird das Ausmaß aktiver Lernzeit, sprich die Zeit, in der sich die einzelnen Schülerinnen und Schüler mit den zu lernenden Inhalten aktiv und engagiert auseinandersetzen, hervorgehoben. Für diese Schlüs-

selfunktion des Unterrichts liegen vergleichsweise robuste Befunde vor. So gilt: Je störungsärmer der Unterricht und je effektiver die Lernzeit, umso höher sind die Leistungen der Schülerinnen und Schüler (Baumert et al., 2004). Aktuellere Ansätze sehen als entscheidend nicht den reaktiven Umgang mit Störungen im Unterricht an, sondern die Prävention von Unterbrechungen des Unterrichtsflusses (Kunter & Voss, 2011). Hierbei wird auch die Rolle der Lehrkraft als Gestalterin der sozialen Verhaltenserwartungen immer wieder betont. Evertson und Weinstein (2013) konnten dabei in empirischen Studien die Wirksamkeit der Etablierung von Regelsystemen und Prozeduren im Unterricht nachweisen. Kounin (1970) prägte hierzu den Begriff *with-it-ness*, also eine Allgegenwärtigkeit der Lehrkraft, welche in der Lage ist, alle im Klassenzimmer ablaufenden Prozesse stets zu überblicken und präventiv einzugreifen.

Klarheit und Strukturiertheit

Bereits in dem von Slavin (1994) entwickeltem QAIT-Modell (Q=Quality, A=Appropriateness, I=Incentives to learn, T=Time) bilden Klarheit, Strukturiertheit und Verständlichkeit die Leitkonzepte der Qualität des Unterrichts. Ebenso setzt Meyer (2003) das Merkmal „Klare Strukturierung“ an den Anfang seiner Liste. Er versteht darunter die Stimmigkeit von Zielen, Inhalten und Methoden, die Folgerichtigkeit des methodischen Gangs, die Angemessenheit des methodischen Grundrhythmus und Regel- und Rollenklarheit. Helmke (2007a) ordnet dem Begriff Klarheit vier Komponenten zu: die akustische Verstehbarkeit, die sprachliche Prägnanz, die inhaltliche Kohärenz und die fachliche Korrektheit. Dabei sind Verständlichkeit und Verstehbarkeit abhängig von Merkmalen der Sprache, wie Lautstärke, Sprechgeschwindigkeit, Pausen, Artikulation, Modulation, Timbre, Gestik und Mimik, und der Verwendung von Hochsprache oder Dialekt. Der Begriff Strukturiertheit hat für ihn zwei Bedeutungen: Zum einen umfasst er alle Merkmale des Informationsangebots, welche das Ziel haben, den Aufbau einer gut organisierten Wissensbasis zu stützen. Gedächtnispsychologisch bedeutet dies unter anderem die Mitteilung der Unterrichts- und Lernziele, transparente Leistungserwartungen, ausdrückliche Verknüpfung der neu vermittelten Informationen mit Vorwissen, das Anbieten von Lernhilfen, aber auch eine gute Dosierung strukturierender Lernhilfen. Zum anderen ist die Strukturiertheit didaktisch gesehen eine Planung und Sequenzierung des Unterrichts. Weitere Kriterien können auch die Verwendung verschiedener Lehr-Lern-Medien und eine Passung der verbalen und schriftlichen Äußerungen, wie zum Beispiel der Tafelan-schrieb, sein.

Aktivierung

Unterricht wird aus einer konstruktivistischen Sicht als erfolgreich bezeichnet, wenn das Verständnis der Schülerinnen und Schüler für fachliche Inhalte gefördert wird. Dies ist Teil des theoretischen Modells der COACTIV-Studie (Baumert & Kunter, 2011). Ebenso betrifft es auch die Lernumgebung, welche so gestaltet sein sollte, dass Lernende zum tiefen Nachdenken bzw. zu einer aktiven mentalen Auseinandersetzung angeregt werden. Leuders und Holzäpfel (2011) fordern dazu den Einbezug aller Bedürfnisse der Schülerinnen und Schüler, soweit möglich. Dies betrifft unter anderem Arbeitsaufträge, welche kognitiv herausfordernd, aber nicht überfordernd gestaltet sein sollen. Ebenso verweisen sie auf die Korrelation des Ausmaßes an kognitiver Aktivierung mit der Nutzung der zur Verfügung stehenden Lernzeit. Dabei fällt dem Öfteren der Begriff des Potentials zur kog-

nitiven Aktivierung (Baumert et al., 2004; Lipowsky, 2015). Lernpsychologisch betrachtet ist dies die Veränderung, Erweiterung, Vernetzung, Umstrukturierung oder Neubildung von bestehenden Wissensstrukturen (Cobb, 1994; Kunter & Voss, 2011). Hierbei geht es ausschließlich um die mentale Aktivität (Baumert & Kunter, 2011). Jedoch fallen nicht nur strukturierende Maßnahmen wie etwa die Anpassung der Lernsituation an die Anforderungen der Lernenden unter den Begriff Aktivierung (Reiser, 2004), sondern auch die Qualität der Beziehung zwischen Lehrkraft und Lernenden. Dies wird unter dem Begriff Unterrichtsklima eingeordnet (Den Brok, Brekelmans & Wubbels, 2004; Fraser, 1991), welcher aber als eines der unschärfsten Konstrukte der Unterrichtsforschung gilt (Gruehn, 2000; Lipowsky, 2015). Um dies zu konkretisieren, wurden Aspekte der Qualität der Beziehung zwischen Lehrkraft und Lernenden in der COACTIV-Studie als emotionale und motivationale Unterstützung der Lernenden durch die Lehrkraft definiert. Dazu zählen unter anderem der respektvolle und geduldige Umgang mit Schülerinnen und Schülern oder eine hohe Sozialorientierung der Lehrkräfte. Letzteres bedeutet hierbei Ansprechbarkeit bei persönlichen oder sozialen Schwierigkeiten (Kunter & Voss, 2011). Ein Zusammenhang zwischen dem Maß an kognitiver Aktivierung und dem kognitiven Lernzuwachs der Lernenden konnte sowohl in der COACTIV-Studie (Baumert & Kunter, 2011) als auch in der QuIP-Studie (Borowski, Olszewski & Fischer, 2010) belegt werden.

Umgang mit Heterogenität

Die Ergebnisse der internationalen Leistungsvergleichsstudie PISA zeigen, dass es in Ländern wie Deutschland nicht ausreichend gelingt, alle Lernenden entsprechend ihrer Voraussetzungen zu fördern (Prenzel et al., 2004). Im Umgang mit heterogenen Leistungsvoraussetzungen im Unterricht können verschiedene Faktoren hilfreich sein (Beck et al., 2008). Diese umfassen zum Beispiel ein Repertoire von Methoden zur Differenzierung, Sachkompetenz und eine damit verbundene diagnostische Kompetenz. Helmke (2007a) nennt sieben Voraussetzungen: den Einstellungswandel hin zu einer positiven Haltung gegenüber Differenzierung, eine diagnostische Kompetenz, Professionswissen und didaktische Expertise, Lehr- und Diagnosematerial, Einbezug außerschulischer Faktoren, Individualisierung und Einhaltung von Standards und Ressourcen an Zeit und Materialien, also auch die nötigen finanziellen Mittel. Möglichkeiten zur Differenzierung im Unterricht werden in der Literatur unter verschiedene Begrifflichkeiten gefasst. Bönsch (1995) und Geppert und Preuss (1978) gebrauchen die Begriffe Binnendifferenzierung oder innere Differenzierung, Corno und Snow (1986) verwenden adaptiver Unterricht bzw. adaptive Teaching und Guldemann, Bischoff und Brühwiler (2005) sprechen von adaptiver Lehrkompetenz. Differenzierungsmaßnahmen können hierbei sowohl den zu vermittelnden sachlichen Inhalt, also die Lernangebote, betreffen als auch die Art und Weise der Aneignung oder die dabei genutzte Sozialform (Hugener, Krammer & Pauli, 2008). Auch das Ausmaß kann differenzieren, von z. B. eher traditionellem Unterricht bis hin zu erweiterten Lehr- und Lernformen oder offenem Unterricht. Als weitere Maßnahmen nennen Hugener et al. (2008) selbstständige Schülerarbeitsphasen, Tempodifferenzierungen, Schwierigkeitsgrad-differenzierung, Kooperationsmöglichkeiten und die Selbstkontrolle der eigenen Resultate. Empirische Studien zum Thema Umgang mit Heterogenität im Unterricht zeigen, dass die Maßnahmen zur inneren Differenzierung eher wenig genutzt werden (Seidel et al., 2006). Positiver von Experten eingeschätzt wird die Öffnung des Unterrichts, was bei Lernenden zu einem höheren Wohlbefinden im Unterricht beitragen kann (Pauli, Reusser, Waldis & Grob, 2003).

Angebotsvariation

Wiechmann (2010, S. 21) fordert eine Öffnung der Sichtweise auf die Methodenvielfalt, denn ein „ausschließlich belehrender Unterricht ist in der Schulrealität ebenso wenig denkbar wie ein rein entdeckender, ein völlig gelenkter Unterricht ebenso unrealistisch wie das vollkommen autonome Lernen. Die Realität eines effektiven Unterrichts liegt zwischen den vier Eckpunkten des Methodenrepertoires. Die didaktisch begründete Wahl der jeweils besten Unterrichtsmethode erfordert eine Kenntnis der spezifischen Leistungsfähigkeit der verschiedenen Unterrichtsmethoden.“ Diese Eckpunkte veranschaulicht er mithilfe folgender Abbildung 5:

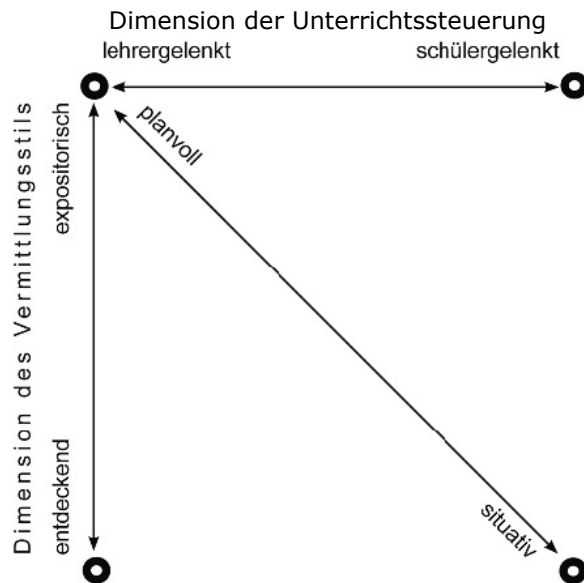


Abbildung 5: Das Entscheidungsfeld der Unterrichtsmethoden (Wiechmann, 2010, S. 19)

So reicht die Dimension der Unterrichtssteuerung von lehrergelenkt bis hin zu schülergelenkt. Die Dimension Vermittlungsstil umfasst expositorisch bis hin zu entdeckend. Ebenso beschreibt Wiechmann (2010) eine Vielzahl an Unterrichtsmethoden, wie Frontalunterricht, Gruppenpuzzle, Stationenarbeit, Wochenplanarbeit, Lernen in Inszenierungen, Lernstückunterricht, Entdeckendes Lernen, Fallstudie, Werkstattarbeit und Projektmethode, welche an unterschiedlichen Positionen im obigen Schema einzuordnen sind. Die Methodenvielfalt ist nach Helmke (2009) die Basis der Angebotsvariation, jedoch besteht ein Unterricht, welcher der Heterogenität der Lernvoraussetzungen, des Stoffs und der zu erreichenden Bildungsziele gerecht wird, aus einer Vielfalt unterrichtlicher Angebote. Er bezieht dies auf eine passende Auswahl an Medien, an Aufgaben bzw. Textsorten, Ansprachen und Lautstärke stimmlicher Äußerungen, an Lernorten und an abwechselnden Lern- und Entspannungsphasen. Zusammenfassend ist anzumerken, dass „der erfolgreichste Unterricht selten derjenige mit einem Maximum an Methodenvariation, sondern der mit einem Optimum“ ist (Helmke, 2009, S. 265). Dieses Optimum ist jedoch von unterschiedlichen Faktoren abhängig und so nur schwer erlernbar. Die Klassengröße bzw. die Zusammensetzung an verschiedenen Lernvoraussetzungen, der Unterrichtsstoff und die gekonnte Umsetzung der gewählten Methode sind Herausforderungen, welche es hierbei zu bewältigen gilt. So resümieren Weinert und Helmke (1996, S. 41): „Instruktion bleibt vermutlich die wissenschaftlich zwar fundierte, aber nur durch gesunden Menschenverstand, praktische Vernunft und plausible Erfahrungsgeneralisierung nutzbare Anwendung von Prinzipien.“

Kooperatives Experimentieren in Kleingruppen

Slavin (1989) definiert kooperatives Lernen als ein Zusammenarbeiten von Schülerinnen und Schülern in kleinen Gruppen, mit dem Ziel, sich gegenseitig beim Lernen des Stoffes zu helfen. Rumann (2005) identifizierte aus verschiedenen Definitionen drei Kennzeichen kooperativen Arbeitens:

- eine Zusammenarbeit in kleinen Gruppen mit einer maximalen Obergrenze von fünf bis sechs Lernenden, also klein genug, um sicherzustellen, dass jedes Mitglied an der Kommunikation beteiligt ist
- gemeinsam zu lösende Aufgaben bzw. gemeinsam zu erreichende Ziele
- eine direkte und unmittelbare Supervision des Lehrers sollte ausgeschlossen sein

Die Erwartungshaltung an diese Art des Lernens ist groß. Ist im Gegensatz dazu die direkte Instruktion im Chemieunterricht als hinderlich für Schülerinnen und Schüler bekannt, Transferwissen aufzubauen, das Interesse für Chemie zu bewahren und die Entwicklung der Lernmotivation zu erreichen (Sumfleth, Rumann & Nicolai, 2004), so gilt kooperatives Lernen als besser geeignet, um Zusammenhänge zu verstehen und komplexe Probleme zu lösen (Rumann, 2005). Rumann (2005) unterscheidet dabei die Begriffe Gruppenarbeit und kooperatives Arbeiten. Letzteres wäre demzufolge eine Unterrichtsmethode, welche eine spezifische Unterrichtsstruktur bedingt. Die meisten Formen sind an die Gruppenarbeit gebunden, jedoch ist es auch möglich, in Partnerarbeit kooperativ zu arbeiten. Um die Effektivität dieser Unterrichtsmethode sicherzustellen, ist es wichtig, auftretende Schwierigkeiten zu kennen. So erwähnt bereits Slavin (1996) den sogenannten free-rider-Effekt. Hierbei beteiligen sich einige Gruppenmitglieder zu Lasten anderer nicht oder nur sehr wenig an der Gruppenarbeit. Dies kann durch eine Aufgabenspezialisierung der Gruppenmitglieder reduziert werden. Ein weiter bekanntes Phänomen ist der Matthäus-Effekt bzw. der the-rich-get-richer Effekt. Hier übernehmen Schülerinnen und Schüler mit höheren motivationalen und kognitiven Eingangsvoraussetzungen zentrale Aufgaben, da andere Gruppenmitglieder als zu schlecht eingestuft werden, ihren Ansprüchen gerecht zu werden. Dies führt zu einem Schereneffekt, der einer Zunahme der Leistungsvarianz entspricht. Der ganging-up-Effekt lässt das Interesse an der Qualität der Arbeitsergebnisse aller Gruppenmitglieder vermissen und führt zu einer Fokussierung auf den eigenen Beitrag (Neber, 1998). Dies kann dazu führen, dass derjenige Lösungsweg präferiert wird, welcher die geringste Anstrengung verspricht (Rumann, 2005). Rumann (2005) konnte trotz allem zeigen, dass bei offenen Aufgaben ein größerer Lernerfolg zu erwarten ist als bei vergleichbarem Frontalunterricht, besonders wenn die Korrektheit von Teilschritten abgesichert oder mit einer Musterlösung verglichen werden konnte (Walpuski, 2006).

Offenheit der Idee, Planung, Durchführung und Auswertung von Experimenten

Nach Janich (1995, S. 621 f.) ist ein Experiment eine „planmäßige Herbeiführung von (meist variablen) Umständen zum Zweck wissenschaftlicher Beobachtung“. Zur Phasierung des Experimentierprozesses existieren unterschiedliche, aber ähnliche Ansätze. Walpuski (2006) unterscheidet die Phasen Hypothese/Idee finden, Experiment planen und durchführen sowie Schlussfolgerung. Labudde und Adamina (2008) unterteilen in bewusst wahrnehmen, Frage/Problem/Hypothese aufwerfen, Untersuchung planen und durchführen, Daten/Ergebnisse auswerten, Erkenntnisse formulieren und Ergebnisse be-

werten (ggf. Untersuchungsplanung ändern und ggf. Frage/Hypothese/Problem umformulieren). Schreiber, Theyßen und Schecker (2009) nennen Planung, Durchführung und Auswertung. Vergleicht man alle Ansätze, so sind die Phasen (a) Finden einer Frage bzw. eines Problems und Generierung von Hypothesen, (b) Planung und Durchführung und (c) Auswertung im Hinblick auf die Idee bzw. die Hypothese als zentral und ebenso als bewährt einzuschätzen (Sumfleth, Wirth & Thillmann, 2013). Als fundamental didaktisches Prinzip gilt insbesondere das schülerzentrierte hypothesengeleitete Experimentieren (Hofstein & Lunetta, 1982). Der naturwissenschaftliche Erkenntnisgewinnungsprozess (scientific inquiry) wird vom National Research Council wie folgt beschrieben:

„Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results“ (National Research Council, 1996, S. 23)

Dabei hat der Wandel des Experimenteinsatzes hin zum selbstregulierten Experimentieren gezeigt, dass der Inquiry-Ansatz die Haltung der Lernenden wesentlich stärker positiv beeinflusst als reine Bestätigungsexperimente (Kipnis & Hofstein, 2008). Bei richtiger Umsetzung analysierten Singer, Hilton und Schweingruber (2006) diverse Vorteile des Einsatzes von Schülerexperimenten. Neben der Zunahme des Verständnisses für naturwissenschaftliche Problemstellungen konnten die Schülerinnen und Schüler ihre experimentellen Fertigkeiten entwickeln, eine Einsicht für das Wesen bzw. die Natur der Naturwissenschaften erreichen und ihre naturwissenschaftliche Argumentationsweise wie auch ihre kollaborativen Fähigkeiten verbessern. Ebenso konnte eine Zunahme der Motivation für Naturwissenschaften verzeichnet werden. Es bleibt jedoch die Aufgabe der Lehrenden, komplexe Ideen und Erfahrungen so zu ordnen, dass die Schülerinnen und Schüler sich sinnvoll an Laboraktivitäten beteiligen können (Lunetta, Hofstein & Clough, 2007). Zudem ergab sich aus verschiedenen Studien, dass eine Fehlerkorrektur, zum Beispiel in Form von Feedback, die Lernwirksamkeit erhöht. Im Gegensatz hierzu können Strukturierungshilfen allein keine zufriedenstellenden Ergebnisse erzielen, was die Schlussfolgerung zulässt, dass Schülerinnen und Schüler das naturwissenschaftlich-experimentelle Arbeiten üben müssen, mit dem Ziel, ihr eigenes experimentelles Arbeiten zu reflektieren, nach Fehlern zu suchen und ihre Schlussfolgerungen zu überprüfen (Walpuski, 2006).

Zusammenfassend kann festgehalten werden, dass die Unterrichtsqualität immer abhängig ist von einem reflektierten Einsatz der Merkmale guten Unterrichts. Hattie (2008) kam zu dem Ergebnis, dass nicht nur eine äußere Ordnung oder Rahmenbedingungen über Qualität von Unterricht entscheiden, sondern ebenso eine innere Ordnung und die Persönlichkeitsstruktur der Lehrpersonen. Eine Übernahme der Schülerperspektive ist von zentraler Bedeutung, was den Lehrerinnen und Lehrern hohe pädagogische Kompetenzen abverlangt (Bolte, 2013). So fasst Gurney (2007) fünf Faktoren effektiven Unterrichts zusammen:

- „Teacher knowledge, enthusiasm and responsibility for learning.
- Classroom activities that encourage learning.
- Assessment activities that encourage learning through experience.
- Effective feedback that establishes the learning processes in the classroom.

- Effective interaction between the teacher and the students, creating an environment that respects, encourages and stimulates learning through experience“ (Gurney, 2007, S. 91).

Smith (1995, S. 590) schreibt: „All teachers do good things some of the time, and all good teachers do bad things some of the time. The differences among teachers lie not only in the proportions of the good and the bad, but also in their awareness of the effects of what they are doing and their readiness to share this awareness with their students.“ Dieses Wahrnehmen und Bewusstwerden der Auswirkungen ihres Handelns im Unterricht liegt begründet in einer hohen Reflexionskompetenz der Lehrenden. Helmke (2009) sieht darin eine Schlüsselbedingung für Unterrichtserfolg und eines der wichtigsten Merkmale der Lehrperson.

3. Reflexionskompetenz

Das Kapitel Reflexionskompetenz unterteilt sich in vier Unterkapitel. Als Basis werden zunächst in Kapitel 3.1 ausgewählte, für diese Studie entscheidende, Definitionen und Modelle zum Konstrukt Reflexion vorgestellt. Anschließend wird Bezug zum Thema Unterricht genommen (Kapitel 3.2). Hierbei wird zunächst die professionelle Handlungskompetenz von Lehrkräften vorgestellt und mit dem Hauptthema der Reflexionskompetenz in Verbindung gesetzt (Kapitel 3.2.1). Im Folgekapitel 3.2.2 wird geklärt, welche Voraussetzungen eine Lehrkraft mitbringen muss, um wissenschaftliche Reflexionsfähigkeit zu entwickeln. Ebenso werden Möglichkeiten zur Förderung der Reflexionskompetenz (Kapitel 3.3) vorgestellt. Diese umfassen die Punkte Lehr-Lern-Labor (Kapitel 3.3.1), Videografie (Kapitel 3.3.2), Learning Study (Kapitel 3.3.3) und Feedback (Kapitel 3.3.4). Die Entwicklung des Kodiermanuals wird theoretisch begründet in Kapitel 3.4.1. Darin werden zahlreiche Manuale vorgestellt, die Einfluss auf diese Studie hatten. Abgerundet wird dieses Kapitel mit aktuellen empirischen Befunden zum Thema Reflexion (Kapitel 3.4.2).

3.1. Definitionen und Modelle zum Konstrukt Reflexion

3.1.1. Definitionen

Literaturrecherche zum Thema Reflexion lässt erkennen, dass die Fähigkeit zu reflektieren als hilfreich angesehen wird, um eine Lehrtätigkeit professionell ausüben zu können. Jedoch sind nur wenige wirkungsvolle Fördermöglichkeiten untersucht (Copeland, Birmingham, La Cruz & Lewin, 1993). Dies liegt sowohl an der Anwendung verschiedener verwandter Begriffe, als Beispiel: Reflexion und Reflexivität (Brendel, 2017), als auch an unzähligen, stark variierenden Definitionen (Abels, 2011; Aufschnaiter, Fraij & Kost, 2019; Calderhead, 1989; Copeland et al., 1993; Gore, 1987; Nitsche, 2014; Noffke & Brennan, 2005; Wischmann, 2015; Wyss, 2013). So ist die Reflexion nach Brendel (2017) anders zu fördern als die in dieser Studie abzugrenzende Reflexivität. Die fehlende einheitliche Definition ist bis heute ein Problem bei der Umsetzung vieler Studien (Al Riyami, 2015; Collin, Karsenti & Komis, 2013; Rivera, 2017).

Die Entwicklung einer für diese Studie passenden Definition führt über Dewey, der seine Idee des Reflective Thinking wie folgt definiert: „Reflective thinking is an active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusion to which it tends constitutes reflective thought“ (Dewey, 1933, S. 9). Demzufolge ist Reflexion ein bewusstes Nachdenken, mit dem ein bestimmtes Ziel verfolgt wird. Unerwartete Ereignisse oder Situationen werden überdacht, um Handlungsvorschläge zu generieren und auszuführen (Wischmann, 2015). Deweys erste Definition musste sich bei der Übertragung auf den Schulkontext vielen Interpretationen unterziehen (Jay & Johnson, 2002). Valli (1997) übertrug Deweys Inhalte auf ihre Definition einer reflektierenden Lehrperson und beschrieb diese als Mensch, welcher auf vergangene Handlungen zurückschauen und über diese urteilen kann sowie sein Lehrverhalten im Hinblick auf Handwerk, Forschung und ethisches Wissen zu ändern vermag (Valli, 1997, S. 70). Wie unterschiedlich die Begriffsdefinitionen sind, zeigt sich in den verschiedenen Ausarbeitungen der Modelle. So dient nach Kolbs Modell des erfahrungsbasierten Lernens Reflexion unter anderem dazu, auf der Grundlage von Erfahrungen persönliche Theorien weiterzuentwickeln (Kolb, 1984). Argyris und Schön (1999) sahen in dem Begriff ein Zusammenspiel von Denken und Handeln, wodurch sich Verhalten weiterentwickeln lässt (Argyris & Schön, 1999, S. 45). Spätere

Reflexionskonzepte unterscheiden sich nach Hatton und Smith (1995) des Weiteren darin, inwieweit Reflexion als bloßes Nachdenken über eine Handlung betrachtet wird oder bereits eine Entwicklung und ein Ausprobieren einer Lösung inkludiert ist. Dewey spricht deswegen nicht nur von Reflective Thinking, sondern auch von Reflective Action, um den gesamten Problemlösungsprozess abzubilden (Hatton & Smith, 1995). Auch Hullfish und Smith (1961) bringen Reflexion mit der Lösung eines Problems in Verbindung und definieren wie folgt: „Reflection differs from the looser kinds of thinking primarily by virtue of being directed or controlled by a purpose – the solution of a problem“ (Hullfish & Smith, 1961, S. 36). Hatton und Smith (1995) sehen das Ziel einer Reflexion darin, das Verhältnis zwischen ursprünglicher Intention und Geschehenem mit resultierenden Schwierigkeiten besser einschätzen zu können. Das Problemlösen ist bei ihnen keine Bedingung einer Reflexion. Sie definieren: „Reflection is a deliberative thinking about action with a view to its improvement“ (Hatton & Smith, 1995, S. 40). Boude, Keogh und Walker (1985) argumentieren für die Inklusion der Emotionen und definieren: „Reflection in the context of learning is a generic term for those intellectual and affective activities in which individuals engage to explore their experiences in order to lead to new understandings and appreciations“ (Boud et al., 1985, S. 19). Ganz allgemein wird mit Reflexion die Verarbeitungsphase (processing phase) nach einer praktischen Erfahrung beschrieben. In dieser Verarbeitungsphase fangen die Personen diese Erfahrungen wieder ein (recapture), grübeln über sie nach (they mull over it), bevor eine Neubewertung stattfindet (evaluate) (Boud et al., 1985). Zimmermann und Welzel (2008) sehen Reflexion als ein von Natur aus normatives Konzept und definieren in Anlehnung an Schön: Reflexionskompetenz ist „Reflexion-über-und-für-die-pädagogische Handlung und besteht in der Fähigkeit, über eine vergangene pädagogische Situation nachzudenken, von allen Seiten zu beleuchten und zu diskutieren, um sie besser zu verstehen und bewusst aus ihr zu lernen“ (Zimmermann & Welzel, 2008, S. 4). Korthagen, Kessels, Koster, Lagerwerf und Wubbels (2002a) sehen Reflexion ebenfalls als normatives Konzept, geben aber zu bedenken, dass jede normative Äußerung über Reflexion fragwürdig ist, solange sie lediglich eine individuelle Meinung ohne theoretische Fundierung über Bildungsziele wiedergibt. In ihrer Arbeit zur Schulwirklichkeit und Lehrerbildung definieren sie wie folgt: „Reflection ist der mentale Prozess zu versuchen, eine Erfahrung, ein Problem oder bestehendes Wissen oder Einsichten zu strukturieren oder zu restrukturieren“ (Korthagen et al., 2002a, S. 63). Sie sind der Meinung, dass durch diese Definition die meisten bestehenden Konzeptualisierungen vereint werden können und Unterschiede von sozialpädagogischen Perspektiven aus bestimmt werden. Jay und Johnson (2002) fanden für ihre Studie folgende Definition, bei der eine Perspektivenübernahme inkludiert ist:

„Reflection is a process, both individual and collaborative, involving experience and uncertainty. It is comprised of identifying questions and key elements of a matter that has emerged as significant, then taking one’s thoughts into dialogue with oneself and with others. One evaluates insights gained from that process with reference to:

- (1) additional perspectives,
- (2) one’s own values, experiences, and beliefs, and
- (3) the larger context within which the questions are raised.

Through reflection, one reaches newfound clarity, on which one bases changes in action or disposition. New questions naturally arise, and the process spirals onward.“ (Jay & Johnson, 2002, S. 76)

In dem Buch „Bildungsstandards und Kompetenzmodelle“ (Gehrmann, Hericks & Lüders, 2010) stellen Leonhard et al. ihre Definition von Reflexionskompetenz vor: „Reflexionskompetenz ist die Fähigkeit, in der Vergegenwärtigung typischer Situationen des schulischen Alltags durch aktive Distanzierung eine eigene Bewertung und Haltung sowie Handlungsperspektiven auf der Basis eigener Erfahrung in Auseinandersetzung mit wissenschaftlichen Wissensbeständen argumentativ zu entwickeln und zu artikulieren“ (Leonhard, Nagel, Rihm, Strittmatter-Haubold & Wengert-Richter, 2010, S. 114). Sie unterteilen Reflexionskompetenz in verschiedene Teilkompetenzen, wie zum Beispiel kontextspezifisches Wissen, sachlogische Argumentation, Selbststeuerung und metakognitive Fähigkeiten. Ebenso sehen sie ein Dilemma einerseits in der präzisen Bestimmung und andererseits darin, Reflexionskompetenz als „Schlüsselqualifikation“ aufzufassen: Leonhard et al. (2010) beschreiben ihre Definition als pragmatische Folge des Wunsches, Reflexionsfähigkeit messen zu wollen. Brendel (2017) unterscheidet zwischen den Begriffen Reflexion und Reflexivität und definiert wie folgt: „**Reflexion** bezeichnet ein Nachdenken über Unterrichtsinhalte, Überzeugungen oder Vorstellungen und geht zurück auf das Konzept des reflexiven Denkens nach Dewey (1997). **Reflexivität** bezeichnet ein Nachdenken über eigene Handlungen oder Handlungsalternativen und fußt auf dem Konzept der reflexiven Praxis nach Schön (1983)“ (Brendel, 2017, S. 19, Hervorhebung im Original). Sie gebraucht in ihrer Arbeit die Begriffe „reflexives Denken“ und „Reflexion“ synonym und ebenso die Begriffe „Reflexivität“ und „reflexive Praxis“. Als Oberbegriff für beide Definitionen verwendet sie „Reflexionsprozesse“.

Trotz diverser Unterschiede der theoretischen Ansätze und Modelle im Detail ist eine Gemeinsamkeit zu erkennen: Reflexion als zentrale Auswirkung auf den Lernprozess (Nitsche, 2014). Die verschiedenen Ansätze stellen eine wichtige Grundlage für die eigenen Untersuchungen dar. Abels (2011) schlägt vor, je nach Kontext und Zielstellung des Reflexionsanlasses eine eigene begriffliche Repräsentation des Konzepts festzulegen, um die fehlende Definition im Bildungskontext auszugleichen. Diese Arbeit übernimmt den Definitionsansatz von Wyss (2013, S. 55), basierend auf den Ausführungen von Schön (1984), Zeichner und Liston (1996) und Jay und Johnson (2002), welcher die wesentlichen Punkte dieser Studie berücksichtigt:

„Reflexion ist ein gezieltes Nachdenken über bestimmte Handlungen oder Geschehnisse im Berufsalltag. Individuell oder im Austausch mit anderen Personen werden die Handlungen oder Geschehnisse systematisch und kriteriengeleitet erkundet und geklärt. Dies geschieht unter Einbezug von:

- (1) Erweitertem Blickwinkel,*
- (2) Eigenen Werten, Erfahrungen, Überzeugungen,*
- (3) Größerem Kontext (theoretische, ethisch-moralische, gesellschaftliche Aspekte).*

Aus dem Prozess werden begründete Konsequenzen für das weitere Handeln abgeleitet und in der Praxis umgesetzt.“ (Wyss, 2013, S. 55)

Die Übertragung dieser Definition auf die hier präsentierte Arbeit kann wie folgt erklärt werden:

Das „gezielte Nachdenken über bestimmte Handlungen oder Geschehnisse im Berufsalltag“ findet Umsetzung in Form von Schülertagen mit anschließender schriftlicher Reflexion der gehaltenen Stunde. Der Punkt „Individuell oder im Austausch mit anderen Personen“ erfolgt in einer Unterscheidung zweier Interventionsgruppen in Selbstreflexion und Reflexion mit vorherigem Feedback. „Systematische und kriteriengeleitete Erkundung und Klärung“ findet Anwendung bei einer Theorieinputphase am Anfang des Semesters und der Bereitstellung der Bibliothek sowie des Internets zur Theorierecherche während der Reflexionen. Die schriftlichen Ausführungen enthalten den „erweiterten Blickwinkel der Studierenden unter Einbezug von eigenen Werten, Erfahrungen und persönlichen Überzeugungen“. Ebenfalls untersucht wird ein Perspektivwechsel, welcher „Theorieperspektive, Fachperspektive, Schülerperspektive und die Politisch/Soziale Perspektive“ miteinschließt. Die Studierenden finden „Begründungen und Konsequenzen“ aus ihrem Handeln, welche auch Inhalte des entwickelten Kodiermanuals darstellen, und verbessern ihre Stunden, um diese im Anschluss erneut zu erproben, sprich „in die Praxis umsetzen“.

3.1.2. Entwicklung der Modelle zum Konstrukt Reflexion

John Dewey

Theorie/Modell:

Als Begründer des Reflexionskonzepts im pädagogischen Bereich wird John Dewey angesehen (Hatton & Smith, 1995), dessen Ansichten und Ideen in fast jeder wissenschaftlichen Arbeit zum Thema Science Education zu finden sind (Wong, Pugh & Dewey Ideas Group at Michigan State University, 2001). Champagne und Klopfer (1977) vermuten sogar, dass jede Theorie über Reflexion und Problemlösung auf die von Dewey entwickelten Grundlagen zurückgeht. In seinem Werk „How we think“ (Dewey, 1933) differenziert er zwischen drei Formen des Denkens: imagination, beliefs und reflective thinking (Dewey, 1933). Das Ziel reflektierenden Denkens besteht nach Dewey nicht darin, Wissen zu erlangen. Vielmehr soll etwas, das für wahr gehalten wird (beliefs), regelmäßig und sorgfältig geprüft werden, und zwar gestützt auf Gründe, die der Ansicht entsprechen (Dewey, 1933). Dabei unterscheidet man fünf Schritte: „(i) a felt difficulty; (ii) its location and definition; (iii) suggestion of possible solution; (iv) development by reasoning of the bearings of the suggestion; (v) further observation and experiment leading to its acceptance or rejection; that is, the conclusion of belief or disbelief“ (Dewey, 1910, S. 72).

Diese problemorientierte Reflexion wird als ein Prozess betrachtet, in dem ein Individuum mit einem Problem konfrontiert wird, welches zu einer tiefen Reflexion führt. Dies kann dazu führen, zugrunde gelegte Annahmen zu hinterfragen und somit ein neues Verständnis einer Situation zur Folge haben (Jungmeister, 2016).

Kritik:

Nichtsdestotrotz bemängeln Champagne und Klopfer (1977) in ihrer Zusammenfassung Deweys Theorie bezüglich der generellen Umsetzung im Klassenzimmer. Sie finden dabei zahlreiche Kritikpunkte, wie zum Beispiel die Schwierigkeit, abstrakte Philosophie auf die Pädagogik zu übertragen oder das hochgesteckte Ziel, Studierende wie Wissenschaftler denken und handeln zu lassen. Auch Rodgers (2002) argumentiert, dass der

Grund für die geringe Umsetzung des Reflexionsmodells nach Dewey in der philosophischen Auslegung mit den über 30 verschiedenen Fachausdrücken besteht, die zur Beschreibung der Komplexität erforderlich sind. Andere kritisierten, dass Dewey die emotionalen Prozesse in seinem Reflexionsprozess vernachlässigt (Boud et al., 1985; Hullfish & Smith, 1961) und entwickelten sein Modell, wie im Folgenden dargestellt, weiter.

H. Gordon Hullfish und Philip Smith

Theorie/Modell:

Hullfish und Smith (1961) erwähnten als weiteren Aspekt die emotionale Ebene als Grundvoraussetzung qualitativ hochwertiger Reflexionen. Sie betonten, dass die Fähigkeit, überhaupt reflektieren zu können, von einem ausgeprägten Vorstellungs-, Gedächtnis- und Empfindungsvermögen abhängt (Moon, 2013) und entwickelten ein leicht verändertes Modell, basierend auf vier Phasen:

- „1. The presence (and recognition) of a problem situation.
 2. Clarification of the problem.
 3. Hypotheses formed, tested and modified.
 4. Action taken on the basis of the best-supported hypothesis.“
- (Hullfish & Smith, 1961, S. 43 f.)

Somit beruht die erste Stufe des Reflexionsmodells von Hullfish und Smith (1961) auf einem Erkennen (recognition), basierend auf einer affektiven Ebene (affective level) (Shiel & Jones, 2003, S. 8). Die letzte Phase des Modells umfasst bereits die Umsetzung der Reflexion in der Handlung (Wyss, 2013).

Kritik:

Das von Hullfish und Smith (1961) entwickelte Konzept des reflexiven Denkens (Reflective Thinking) scheint bereits das im Folgenden erklärte Konzept der Reflexion in der Aktion (reflection-in-action), entwickelt von Schön (1984), zu antizipieren, obwohl es für viele fraglich ist, diese Form überhaupt als Reflexion zu bezeichnen (Moon, 2013). Im Gegensatz zu Schön (1984) fehlt eine zeitliche Abgrenzung der verschiedenen Reflexionsarten (Moon, 2013; Wyss, 2013).

Donald Schön

Theorie/Modell:

Donald Schön ist einer der bedeutendsten Vertreter der Theorie des reflektierten professionellen Lernens (reflective professional learning). Er bezieht sich dabei auf die grundlegenden Arbeiten von Dewey (1910, 1933) und später auch von David Kolb (1984). Sein bekanntester Ansatz ist die Unterscheidung zwischen Reflexion in der Aktion (reflection-in-action) und Reflexion nach der Aktion (reflection-on-action) aus seinem Werk „The Reflective Practitioner“ (1984), welches zudem den zeitliche Aspekt von Reflexionen berücksichtigt. Ersteres bezeichnet eine unmittelbare Bewertung der Aktionen noch während der Aktion selbst (Hilzensauer, 2008). Der Prozess verläuft in sechs Phasen:

reflection-in-action

1. Eine problematische Situation entsteht.
2. Es erfolgt eine Problemdefinition.
3. Es erfolgt die Verwirklichung der ersten Problemdefinition.
4. Es erfolgt eine gleichzeitige experimentelle Prüfung der ersten Problemdefinition.
5. Die hypothetische Definition wird der gegebenen Situation übergestülpt.
6. Bewertung der Problemdefinition in Bezug auf die Ergebnisse des Experiments.

Der zweite Begriff, also *reflection-on-action*, wird als Reflektieren nach einer Situation beschrieben. Die Schritte ähneln den bereits genannten mit dem Unterschied, dass handelnde Personen ihre Erfahrungen nachträglich dokumentieren und erst dann über Erfolgs- und Misserfolgskriterien reflektieren.

Kritik:

Obwohl Schöns Arbeit viel Aufmerksamkeit in der universitären Lehramtsausbildung erhalten hat, ist, wie bei Hullfish und Smith bereits angemerkt, die Abstammung nicht immer genau nachzuvollziehen (Munby, 1989). Die zahlreichen unterschiedlichen Auslegungen könnten auf Schwierigkeiten bzw. Unklarheiten bezüglich Schöns Konzept zurückzuführen sein. Andere verweisen wiederum darauf, dass dies Ursache der Unvollständigkeit des Konzepts, bezogen auf fehlende Perspektivenübernahmen oder emotionale Prozesse, wäre (Munby, 1989). Ebenso wird das Konzept *reflection-in-action* als Mitdenken und *reflection-on-action* als – das tendenziell eher mit Reflexion in Verbindung gebrachte – Nachdenken verortet (Straßer, 2008). Umstritten am Konzept *reflection-in-action* ist, dass „es eine hohe Kompetenz abverlangt, parallel zur Handlung zu reflektieren und augenblicklich Veränderungen einzuleiten“ (Abels, 2011, S. 53). Ein weiterer Kritikpunkt ist der fehlende gesellschaftliche Aspekt (Wyss, 2013). Schön (1983) beschränkt sich darauf, Reflexion als individuellen Prozess einer Person zu sehen. Ostermann und Kosskamp (2004) betonen jedoch die Wichtigkeit gemeinsamer Reflexion im sozialen Kontext. Auch finden sich laut Hilzensauer (2008) in Schöns Ausführungen kaum Hinweise auf das Reflektieren über den Lernprozess selbst. Yost, Sentner und Forlenza-Bailey (2000) nennen als weitere Reflexionsform die Reflexion für die Aktion (*reflection-for-action*): Sie definieren Reflexion im Rahmen der Planung des Unterrichts, sprich vor dessen Durchführung unter Bezug auf vorangegangene Praxis. Diese Form beschreiben sie als gewünschte Folge der anderen beiden Formen der Reflexion. Trotz dieser diversen Kritikpunkte ist Schöns Vorstellung zu Reflexion und Professionalisierungsprozessen eine der einflussreichsten Arbeiten, welche in viele Diskussionen über die Aus- und Weiterbildung von Lehrpersonen immer wieder einfließt (Wyss, 2013).

David A. Kolb

Theorie/Modell:

Kolb gilt als (Mit-) Begründer des sog. Erfahrungslernens (*experiential learning*) (Hilzensauer, 2008) und baut dabei unter anderem auf den Arbeiten von John Dewey auf. Eines der einflussreichsten Lehr-/Lernmodelle bildet dabei der Lernkreislauf (*learning cycle*) (Kolb, 1984) (Abbildung 6). Bestehend aus vier Phasen ist die Ausgangslage hier

eine konkrete Erfahrung. Gefördert werden der Prozess, Erfahrungen machen zu können und diese zu reflektieren sowie die Fähigkeit, Beobachtungen abstrahieren und Schlüsse ziehen zu können, und ebenfalls die Fähigkeit, Erkenntnisse in praktische Situationen zu überführen und diese zu reflektieren.

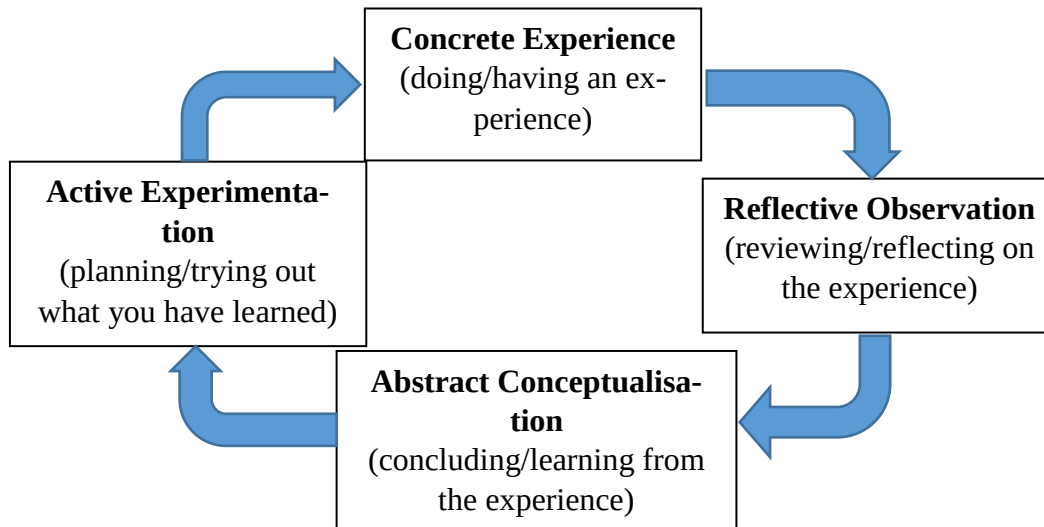


Abbildung 6: Learning cycle (Kolb, 2014, S. 21)

Kritik:

Obwohl eines der am häufigsten genutzten Modelle, ist der Lernkreislauf von Kolb auch eines der am meist kritisiertesten (Greenaway, 2008). Um die Modellkritik auf der Ebene der Reflexion zu belassen, ist hier anzumerken, dass Kolb Lücken in der Beschreibung der einzelnen Prozesse der Reflexion aufweist; diese stehen lediglich in Beziehung zu den drei anderen Teilen des Kreislaufs (Boud et al., 1985). Unbestritten ist die Bedeutung der Reflexion für das Lernen, kritisiert wird jedoch die Möglichkeit der Übertragung auf unterschiedliche Lernstile (Boud et al., 1985; Jarvis, 1995), welche nur schwer denkbar scheint.

Hilzensauer (2008) bemerkt zusätzlich, dass auch soziale und kulturelle Einflussfaktoren weitgehend unbeachtet bleiben, ebenso wie eine detaillierte Erklärung über den Zusammenhang von Lernen und dem Aufbau von Wissen. Den Erfolg des Modells begründet er mit der Anschaulichkeit und vor allem der Handhabbarkeit als eine praktische Grundlage für Lehrer und Lehrerinnen.

David Boud, Rosemarie Keogh & David Walker

Theorie/Modell:

1985 ergänzten Boud et al. bei ihrem Modell die bei Kolb fehlende Erklärung über interne (Gedanken-)Prozesse beim Reflektieren. Basierend auf Dewey und Kolb gehen sie von einer vorangegangenen Erfahrung aus, welche als zentrale Voraussetzung für das Lernen angenommen wird. Jungmeister (2016) erkennt sieben Elemente des Reflexionsprozesses in Bezug auf selbstgesteuertes Lernen (Boud et al., 1985, S. 20 ff.):

1. Rückkehr zu bzw. Integration von direkter Erfahrung
2. Einbezug der Gefühle

3. Assoziationen
4. Integration der Assoziationen
5. Validierung der Befunde
6. Bewertung der Befunde
7. Festlegungen von Beschlüssen bzw. Maßnahmen

Boud et al. (1985) argumentieren für den Einbezug der Emotionen in den Reflexionsprozess, da sie als Katalysator, als Quelle und als Nebenprodukt auftreten können, und ergänzen diese im Vergleich zu den bisherigen Modellen. Reflexionen – basierend auf Emotionen – und Emotionen können einen hohen Nutzen im Lernprozess haben (Mc Manus, 2011).

Kritik:

Jungmeister (2016) gibt zu bedenken, dass in diesem Modell kein Hinweis auf die Reflexion über das Lernen selbst zu finden ist.

Max van Manen (1991)

Theorie/Modell:

Bereits 1977 erarbeite Van Manen ein Modell (Tabelle 2), welches auf den drei Stufen technische Rationalität (technical rationality (TR)), praktisches Handeln (practical action (PA)) und kritische Reflexion (critical reflection (CR)) aufbaute (Ballard, 2006). Während die erste Ebene (TR) Effektivität und Effizienz zum Erreichen eines Ziels auszeichnet, ohne für Kritik offen zu sein, erreicht die zweite Ebene (PA) bereits ein Hinterfragen der Ziele und der verwendeten Mittel, mit denen diese erreicht werden können (Shiel & Jones, 2003). Die kritische Reflexionsebene beinhaltet zusätzlich moralische und ethische Aspekte und verortet Analysen in einen weiten sozio-historischen Kontext (Hatton & Smith, 1995). So ist in diesem Modell bereits eine hierarchische Struktur zu erkennen.

Tabelle 2: Ebenen der Reflexion von Van Manen (1977), zitiert nach Ballard (2006)

TR	Consists of responses that deal with the technical application of educational knowledge and basic curriculum principles, such as are the students doing what the teacher asked them to do. At this level, the contexts of the classroom, school, community, and/or society are not taken into consideration.
PA	The teacher becomes more concerned with clarifying assumptions and predispositions while assessing the educational consequences toward which a teaching action leads. S/he analyzes student and teacher behaviors to see if and how goals are met.
CR	At this level, educators are concerned with worth of knowledge and the social circumstances useful to students without personal bias. S/he asks her/himself several questions such as what were the strengths of the lesson, what should be changed, and was the content covered important to the students?

Kritik:

Barnett (1997) kritisierte die fehlende emotionale Ebene. Diese wurde bereits zu einem früheren Zeitpunkt beispielsweise in den Arbeiten von Hullfish und Smith (1961) als essentiell für qualitativ hochwertige Reflexionen erachtet.

Jack Mezirow**Theorie/Modell:**

Mezirow übernahm einen sehr ähnlichen Ansatz wie Van Manen und baute zusätzlich sein Modell auf Deweys Vorbild auf und dessen Erkenntnis, dass Reflexion das Ziel haben sollte, eine verbesserte Erkenntnis bzw. Handlung zu erreichen. Er unterschied als Erster zwischen reflexiven und nicht-reflexiven Aktionen (Jungmeister, 2016) und fand drei Formen von nicht reflektierenden Aktionen (non-reflective action) (Gewohnheit, gedankenvolle Aktion und die Selbstbeobachtung) sowie zwei Formen des reflektiven Handelns (reflective action). Das niedrigere Reflexionslevel wurde dabei in Inhalt und Prozess unterteilt. Die höhere Ebene nannte er Prämissen-Reflexion (premise reflection) (Abbildung 7). Mezirow und Associates unterscheiden folglich zwischen inhaltlicher und prozessorientierter Reflexion und zeigen ebenso verschiedene Qualitäten von Reflexion in einer hierarchischen Stufung auf (Brendel, 2017).

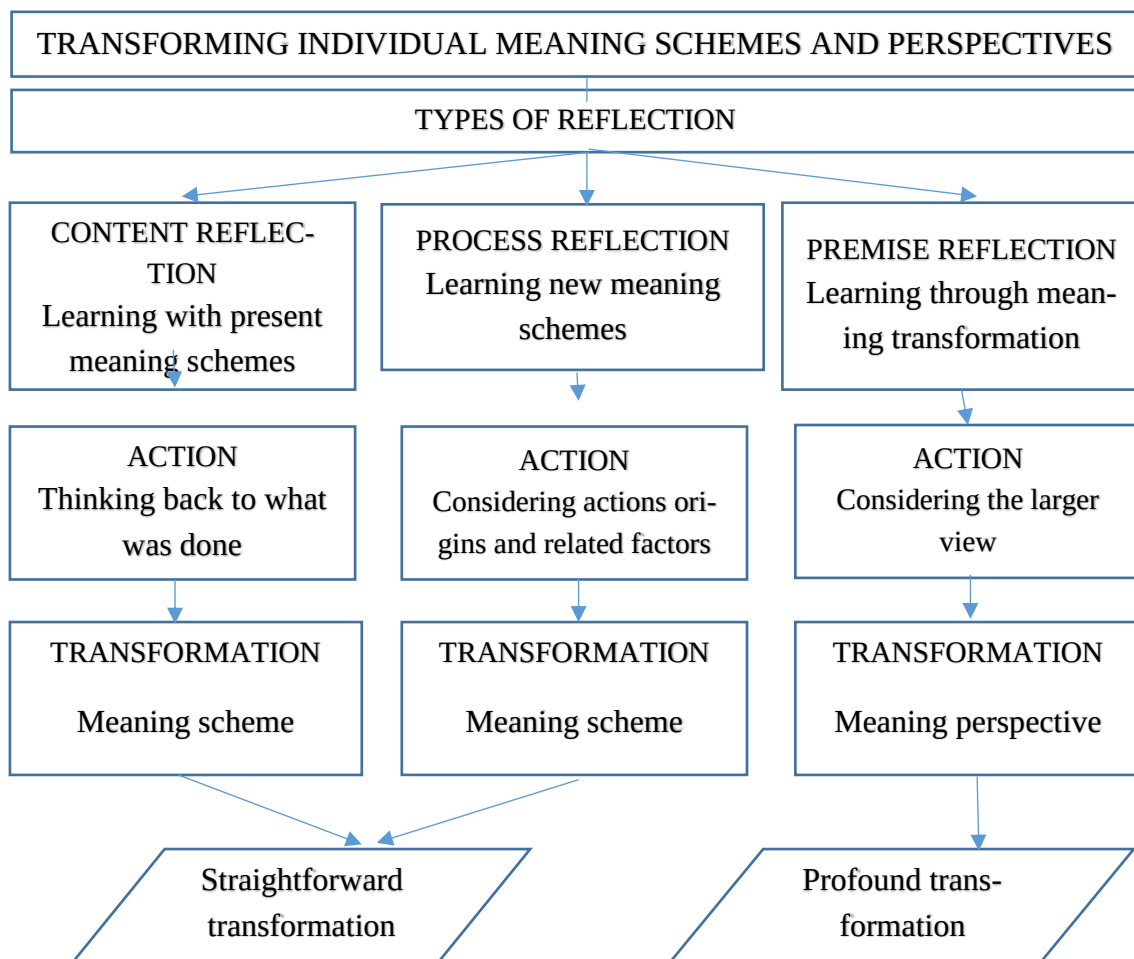


Abbildung 7: Reflexionsmodell nach Mezirow, 1995 (Abbildung übernommen von Kitchenham, 2008, S. 115)

Kritik:

Mezirow (1997) selbst überarbeitete seine Arbeit zur kritischen Reflexion (critical reflection) und präsentierte zwei neue Aspekte. Er ergänzte die kritische Reflexion von Annah-

men (critical reflection of assumptions), bei der nicht nur der Rückblick auf eine Ausgangssituation vollzogen wird, sondern zusätzlich Vermutungen bzw. Voraussetzungen betrachtet werden, welche im Reflexionsprozess involviert waren. Des Weiteren nahm er den Punkt kritische Selbstreflexion von Annahmen (critical self-reflection of assumptions) mit auf, welcher die Kritik an der Voraussetzung, über welcher der Lernende ein Problem definiert, beinhaltet (Kitchenham, 2008; Mezirow, 1998).

Willis D. Copeland

Theorie/Modell:

Copeland et al. (1993) versuchten, ihre Arbeit auf der Natur der Reflexion selbst aufzubauen. Sie stellten sich Fragen wie „Was ist Reflexion in der Praxis?“ und „Wie würde man einen reflective practitioner erkennen, wenn man ihn sieht?“ (Copeland et al., 1993, S. 348). Die Absicht dahinter war, eine Reihe von Kennzeichen reflexiver Praxis herauszuarbeiten. Copeland et al. (1993) sahen Reflexion als Denkprozess an und entwickelten vier Oberstufen: Problemidentifikation, Lösungen generieren, Lösung ausprobieren und Reflektierender Praktiker. Diese sind wiederum wie folgt unterteilt:

Problemidentifikation

1. Ein Problem wird identifiziert.
2. Das Problem stammt aus einer konkreten Situation in der Praxis.
3. Das Problem ist von großer Bedeutung für den Practitioner.
4. Das Problem ist ausschlaggebend für erfolgreiches Unterrichten/Lernen in dem Kontext, in dem es identifiziert wurde.

Lösungen generieren

5. Mögliche Lösungsvorschläge zum Problem werden generiert.
6. Lösungsvorschläge sind theoriebasiert oder gründen auf Theorien, Annahmen oder Forschungsergebnissen, welche explizit vom Practitioner eingehalten und verstanden werden.
7. Das Generieren einer Lösung veranlasst den Lehrer/die Lehrerin zu einer kritischen Betrachtung seiner/ihrer professionellen Handlungsweisen und führt zur gewünschten Handlungsweise.
8. Die angestrebten Lösungen sollen positive Konsequenzen auf den Unterricht haben.

Lösung ausprobieren

9. Ein Lösungsweg wird gewählt.
10. Der gewählte Lösungsweg wird umgesetzt.
11. Die Lösung wird an der Größe des Effektes der Zielaktion und den Konsequenzen des Effekts auf die Schülerergebnisse abgewogen.

Reflektierender Praktiker

12. Der reflexive Prozess führt zu einer Verbesserung des Lehrerverständnisses im professionellen Kontext, in welchem das Problem identifiziert wurde.

(Copeland et al., 1993)

Mithilfe dieser Stufen sollen Forscher die Gedanken der Lehrer im Reflexionsprozess besser analysieren können und so in der Lage sein, die Reflexionskompetenz zu fördern (Copeland et al., 1993, S. 347). Hauptursache dafür sehen sie in den geringen bis fehlenden Möglichkeiten für Lehrer, Rückmeldung über ihren Unterricht nach dem Ende ihrer Ausbildung zu erhalten (Ferri & Blum, 2018).

Kritik:

Cornford (2002) kritisiert die hohe Zahl der Artikel (sie nimmt Bezug auf Copeland et al., 1993; Tom, 1985; Hatton & Smith, 1995; Kenneth M. Zeichner, 1993) zum Thema Reflexion, welche versuchen, die verschiedenen Ansichten zur Natur der reflexiven Unterrichtspraxis in eine ordentliche Taxonomie zu bringen. Sie weist darauf hin, dass diese Versuche nicht die allgemeine Zustimmung getroffen hätten oder eine größere Neigung hervorriefen, die unterschiedliche Effektivität, die verschiedenen Paradigmen bzw. Prozesse bei der Durchführung genauer zu untersuchen.

Fred Korthagen

Theorie/Modell:

Das ALACT-Modell (Abbildung 8) ist selbst ein Prozessmodell und beschreibt nicht die Qualität der Reflexion. Korthagen, Kessels, Koster, Lagerwerf und Wubbels (2002b) beschreiben Lernen als zyklischen Prozess der Reflexion innerhalb von fünf Phasen. Eine konkrete Handlung löst den Prozess aus. Verbunden damit sind ein oder mehrere Handlungsziele. Im Anschluss folgt ein Rückblick auf die Handlung. Auslöser hierbei ist allgemein etwas Unerwartetes, wodurch ein Handlungsziel nicht oder nicht leicht erreicht wurde. In Phase drei wird bestimmten Aspekten eine spezielle Bedeutung beigemessen. Oft findet dies parallel zum Finden von Verbindungen statt. So kann in der nächsten Phase ein anderer Ansatz gewählt werden, der zum Erreichen des Ziels geeigneter scheint. Letztendlich wird in einer neuen, ähnlichen Situation eine neue Handlung unternommen, um das gewünschte Ziel zu erreichen (Korthagen et al., 2002a).

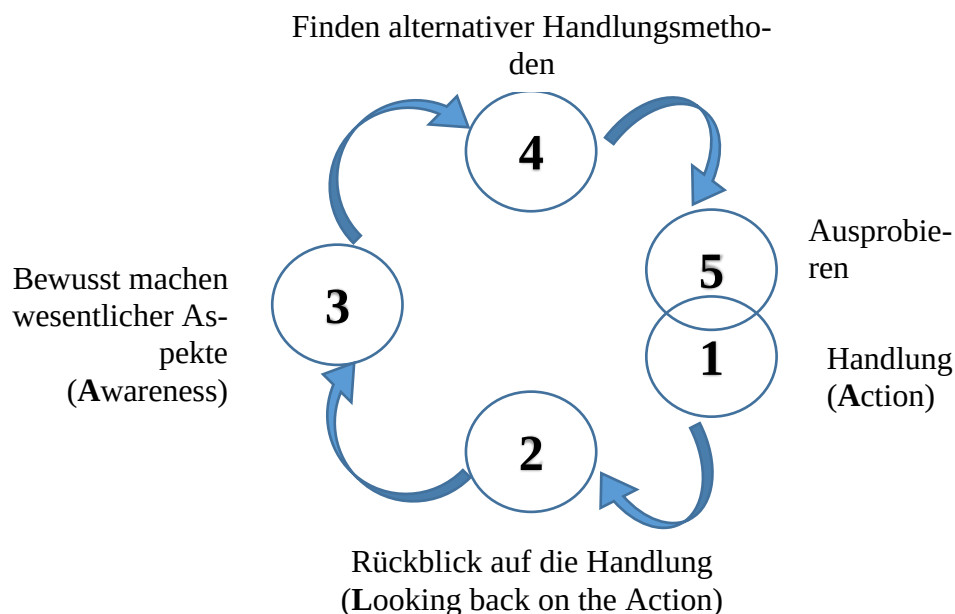


Abbildung 8: ALACT-Modell (Korthagen et al., 2002b, S. 49)

Kritik:

Um die fehlende Qualität der Reflexionen zu ergänzen, entwickelten Korthagen und Vasalos das Zwiebel-Modell (Korthagen & Vasalos, 2009). Es beschreibt sechs Stufen der Reflexionstiefe:

1. „environment,
2. behaviour,
3. competencies,
4. beliefs,
5. professional identity,
6. mission” (Korthagen & Vasalos, 2009, S. 5)

In diesem Modell werden die Ebenen dargestellt, die für das Handeln von Lehrpersonen verantwortlich scheinen und somit großen Einfluss auf die Qualität der Handlungen haben (Wyss, 2013). Direkt beobachten lassen sich hier nur die ersten beiden Ebenen, Umwelt und Verhalten. Schwieriger zu verändern, da über lange Zeit aufgebaut und unsichtbar, sind die restlichen Faktoren: die Kompetenzen, die Überzeugungen einer Lehrperson, die Selbst-überzeugung und der Lebenszweck (Wyss, 2013). Korthagen und Vasalos (2005) sprechen von zentraler Reflexion (core reflection), wenn eine Reflexion auch die Ebenen fünf und sechs einbezieht. Also die höchste Stufe einer Reflexion.

Altrichter & Posch & Spann**Theorie/Modell:**

Die Ziele der Aktionsforschung von Altrichter, Posch und Spann (2018) bestehen darin, gleichzeitig Erkenntnis und Entwicklung anzustreben. Dies geschieht als Ergebnis von Reflexion und Aktion. Im Sinne des Theorie-Praxis-Bezugs sollen sowohl die untersuchte Praxis als auch das praktische und wissenschaftliche Wissen über die Praxis weiterentwickelt werden. Dabei ist ein Merkmal das „In-Beziehung-Setzen von Aktion und Reflexion“ (Altrichter et al., 2018). Der prozesshafte Charakter von Reflexionsformen wird dabei mithilfe eines Zirkels dargestellt, ausgehend von der Aktionsidee über die Aktion und die Reflexion der Aktion zur Rückwirkung auf die (praktische) Theorie, die dann wiederum eine Aktionsidee generiert (Liebsch, 2010) (Abbildung 9). Dabei wird der Aktionsforschungsprozess langfristig angesetzt und als Kreislauf betrachtet, der sich zu einer nach „oben“ führenden Spirale entwickelt.

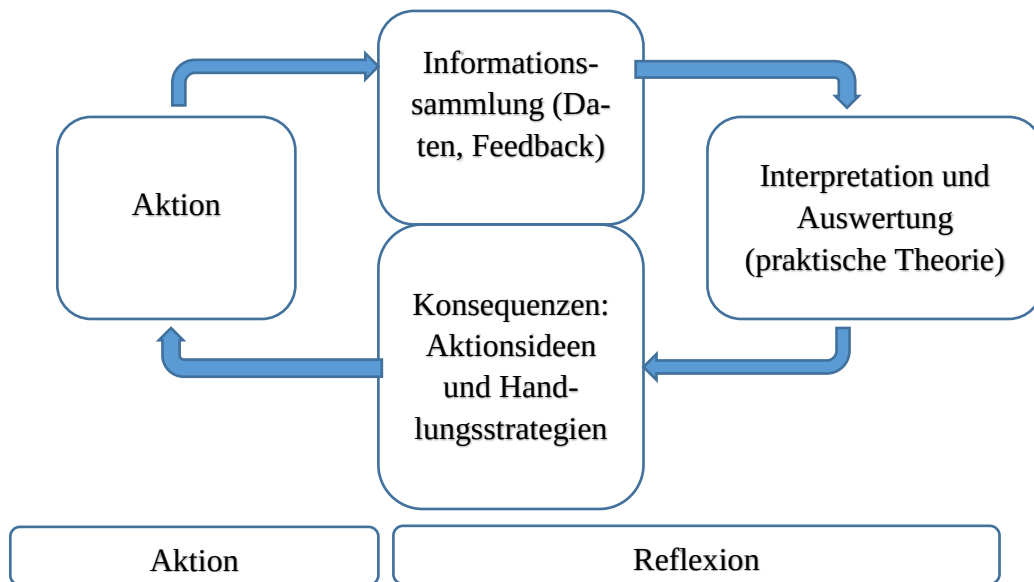


Abbildung 9: Der Kreislauf von Aktion und Reflexion der Aktionsforschung (Altrichter, Posch & Spann, 2018, S. 14)

Kritik:

Kritik besteht in der von Altrichter et al. getroffenen impliziten Annahme, dass die Bewegung innerhalb des Zirkels bereits reflexiv ist. Somit ist unklar, wo und warum das Ausgangsmoment nach einem Durchlauf inhaltlich angereichert wurde (Liebsch, 2010).

Zusammenfassung (Tabelle 3):

Tabelle 3: Zusammenfassung der Modelle zum Konstrukt Reflexion

Vertreter	Ergänzungen zum vorherigen Modell	Kernaussagen/Modell
Dewey (1933)		Lernen setzt Handeln voraus, primäre Erfahrungen werden maßgeblich durch Reflexion zu sekundären Erfahrungen
Hullfish und Smith (1961)	„recognition“ und „affective level“ als Basis der ersten Stufe des Reflexionsmodells	Emotionale Ebene als Grundvoraussetzung
Schön (1983)	Berücksichtigt zusätzlich den zeitlichen Aspekt der Reflexion	Reflection-in-action; reflection-on-action
Kolb (1984)	Ausgangslage hierbei ist eine konkrete Erfahrung	Learning cycle
Boud et al. (1985)	Erklärung über interne (Gedanken-)Prozesse beim Reflektieren	Konkrete Erfahrungen; Reflexion; neue Perspektive
Van Manen (1990)	Ordnet Reflexionen in ein kognitiv-hierarchisches Modell	Grundmodell: drei Stufen, technical rationality (TR), practical action (PA) und critical reflection (CR)

Mezirow	Reflexion sollte das Ziel haben, eine verbesserte Erkenntnis bzw. Handlung zu erreichen	Unterscheidung zwischen Reflexion, nicht-Reflexion und kritischer Reflexion
Copeland (1993)	Stufen des Reflexionsprozesses zur Förderung der Reflexionskompetenz	Gedanken zur Förderung der Reflexionskompetenz
Korthagen et al. (2002b)	Prozessmodell und zusätzliches Modell zur Abschätzung der Qualität von Reflexionen	ALACT-Modell und Zwiebel-Modell
Altrichter et al. (2007)	In-Beziehung-Setzen von Aktion und Reflexion	Kreislauf von Aktion und Reflexion der Aktionsforschung

3.2. Reflexion im Unterricht

3.2.1. Reflexion und Profession

Um genau zu verstehen, was Profession – einerseits bezogen auf den Lehrberuf und zusätzlich bezogen auf die Reflexionskompetenz – genau bedeutet, gilt es zunächst zu klären, wie man den Begriff ableiten und schließlich in Bezug auf das Thema Reflexion definieren kann.

Das lateinische Verb ‚profiteri‘ bedeutet übersetzt ‚erklären‘, oder auch ‚öffentlich bekennen‘ (Roters, 2012). Vor dem Hintergrund des Luhmannschen Gesellschaftsmodells (Luhmann, 1981) interpretiert Stichweh (1992) eine Profession mit den Kriterien Autonomie des Professionellen, das Handeln gegenüber den Klienten, die Förderung gesellschaftlicher Werte, die wissenschaftliche Ausbildung und die Vorgabe von Standards.

Schwendenwein (1990) impliziert bereits die Reflexion mit folgenden sechs Kriterien: berufsethische Reflexion, subjektivitätsbezogene Reflexion, wissenschaftsbezogene Reflexion, praktisch-wissenschaftliche Kompetenz, autobiografische Reflexion und kritische Reflexion. Hierbei sollen professionsspezifische Leitziele gekannt und umgesetzt, die Arbeits- und Handlungsvoraussetzungen diagnostiziert, relevante Theorien angewendet, ein standardisiertes Methodenrepertoire pädagogischer Fähigkeiten und Fertigkeiten aufgewiesen, die persönlichen Erfahrungen einbezogen und die durchgeführten Handlungen bewertet werden.

Die starke institutionelle Eingebundenheit der Berufsgruppe der Lehrkräfte hat teilweise zu einer Etikettierung als Semi-Profession geführt (Combe, 1997; Feiks & Krauß, 2001; Hericks, 2006; Oser, Zutavern & Patry, 1990; Pfadenhauer, 2005; Terhart, 2011). Felder der Medizin und des Rechts werden als klassische Professionen aufgeführt, da diese als freie Berufe erachtet werden (Pfadenhauer, 2005; Terhart, 2011). So wäre nach Pfadenhauer (2005) alles, was zum Beispiel Mediziner und Juristen tun, als professionelles Handeln zu begreifen, da sie eben zuvor als Professionelle definiert worden sind. Ein Kausalschluss, der verlangt, gewisse Standards für eine Professionalisierung festzulegen. Es ist jedoch nicht mehr gebräuchlich, eine strenge Abgrenzung zwischen Beruf und Profession durchzuführen. Das moderne Verständnis fordert eine akademische Ausbildung als Voraussetzung und eine Beschäftigung mit komplexen und insofern immer riskanten

technischen, wirtschaftlichen, sozialen und humanen Problemlagen ihrer Klienten (Terhart, 2011).

Im Hinblick auf die pädagogische Professionalität wird die Professionalisierung des Lehrberufs mit folgenden Argumenten begründet (Daheim, 1992):

- Das Wissen von Lehrerinnen und Lehrern ist generiertes Sonderwissen,
- sie sind sowohl Experten für das Fachwissen als auch für die zu steuernden Prozesse,
- es existiert eine organisatorische Rahmung,
- es besteht keine konventionelle Absicherung und
- eine Reflexivität in Bezug auf die Kontrolle der Prozesse.

Lebenslanges Lernen und beständige Weiterentwicklung sind große Herausforderungen im Lehrberuf. Es ist also nötig, das Verständnis darüber, was eine gute Lehrperson ausmacht, genauer zu charakterisieren bzw. herauszufinden, welche kritischen Punkte der Lehrerentwicklung entgegenwirken (Neophytou, Koutselini & Kyriakides, 2011). Die Rolle der Lehrperson bei der Qualität des Unterrichts ist unumstritten (Hattie, 2008; Helmke, 2009; Lipowsky, 2006; Reusser et al., 2010; Weinert & Helmke, 1996). Jedoch liefern die bisherigen theoretischen und empirischen Befunde ein nicht immer einheitliches Bild (Weinert & Helmke, 1996). Gründe hierfür sind zahlreich, können zum Beispiel an kultur- und milieuspezifischen Überzeugungen liegen oder am historischen Bezug. Zusätzlich werden in den verschiedenen Schularten unterschiedliche Erwartungen an die Lehrpersonen gestellt (Terhart, 2007). Als Kern pädagogischer Professionalität jedoch gilt hierbei das Öfteren Reflexivität, unabhängig vom Professionsverständnis, welches einen berufsbiografischen, handlungstheoretischen, kompetenztheoretischen oder strukturtheroretischen Ansatz haben kann (Denner & Gesenhues, 2013). So erachtet Bosse (2012) die Reflexionskompetenz als Schlüsselkompetenz, über die angehende und vor allem praktizierende Lehrende verfügen müssen (Bold, 2005; Combe & Kolb, 2004; Neuweg, 2005). Reh (2004) versteht die Professionalisierung im Lehrberuf entsprechend der Steigerung der Reflexivität. Professionalisierung in Bezug auf den Lehrberuf ist folglich eng mit der Förderung der Reflexionskompetenz verknüpft (Bennewitz, Rothland & Terhart, 2011). Unter dem Begriff Kompetenz versteht Frey (2006, S. 31) „ein Bündel von körperlichen und geistigen Fähigkeiten [...], die jemand benötigt, um anstehende Aufgaben oder Probleme zielorientiert und verantwortungsvoll zu lösen, die Lösungen zu reflektieren und zu bewerten und das eigene Repertoire an Handlungsmustern weiterzuentwickeln.“ Weinert (2001, S. 27 f.) definiert Kompetenzen als „die bei Individuen verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten, bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können.“ Es stellt sich die Frage, wie sich Reflexionskompetenz in aktuellen Kompetenzmodellen abbilden lässt. Vogelsang (2014) unterscheidet dabei zwei Klassen von Kompetenzmodellen, welche Überschneidungen aufweisen können. Die erste Klasse knüpft an den Begriff Standard an, die zweite Klasse baut auf dem Begriff Professionswissen auf.

Der Begriff Standard wird definiert als „Wissensbestände, die in absolut notwendigerweise angeeignet werden müssen und die hierin auch einem handlungsorientierten Gütemaßstab standhalten“ (Oser, 1997, S. 27). Als Expertise für die ständige Konferenz der

Kultusminister (KMK) entwickelte Terhart (2005) ein einflussreiches Modell zur Beschreibung der Kompetenzen. Diese umfassen sowohl die erste als auch die zweite Phase der Lehramtsausbildung. Für Absolventen der ersten Phase werden die vier Inhaltsbereiche 1. Standards für die Unterrichtsfächer, 2. Standards für die Fachdidaktiken, 3. Standards für das erziehungswissenschaftliche Studium und 4. Standards für die schulpraktischen Studien unterschieden. Eine orthogonale Dimension inkludiert hierbei den Bereich Reflexion, neben den Bereichen Wissen, Urteil und Können. Aus Terharts Vorschlag verabschiedete die KMK 2004 die allgemeinen Standards für die Lehrerbildung mit den vier Kompetenzbereichen Unterrichten, Erziehen, Beurteilen und Innovieren (KMK, 2004). Die weitere Unterteilung in spezifische Kompetenzen beinhaltet immer wieder die Reflexionskompetenz. So sind Lehrerinnen und Lehrer „Fachleute für das Lehren und Lernen. Ihre Kernaufgabe ist die gezielte und nach wissenschaftlichen Erkenntnissen gestaltete Planung, Organisation und Reflexion von Lehr- und Lernprozessen sowie ihre individuelle Bewertung und systemische Evaluation. Die berufliche Qualität von Lehrkräften entscheidet sich an der Qualität ihres Unterrichts“ (KMK, 2004, S. 3). Unter die inhaltlichen Schwerpunkte der Ausbildung fällt die „Begründung und Reflexion von Bildung und Erziehung in institutionellen Prozessen“ (KMK, 2004, S. 4). Zusätzlich wird die Förderung der Entwicklung der Kompetenzen beschrieben, durch

„den Einsatz von Videostudien

die persönliche Erprobung und anschließende Reflexion eines theoretischen Konzepts in schriftlichen Übungen, im Rollenspiel, in simuliertem Unterricht oder in natürlichen Unterrichtssituationen oder an außerschulischen Lernorten

die Analyse und Reflexion der eigenen biographischen Lernerfahrungen mithilfe der theoretischen Konzepte

die Kooperation bei der Planung sowie gegenseitige Hospitation und gemeinsame Reflexion“ (KMK, 2004, S. 6).

Die Ausführung der vier Kompetenzbereiche ist unterteilt in den theoretischen Ausbildungsabschnitt und den praktischen Ausbildungsabschnitt. Dabei fällt auf, dass im theoretischen Ausbildungsabschnitt neben den Operatoren Kennen und Wissen auch Reflektieren von bestimmten Inhalten als Bestandteil der Kompetenz formuliert ist. Diese Standards sind noch allgemeiner, fachunabhängiger Natur und müssen weiter ausdifferenziert werden. Dies ist in einem ersten Ansatz bei den von der KMK (2008) formulierten Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung zu finden. Für eine genauere Ausgestaltung des Curriculums entwickelte (Großebram, 2013) anhand einer qualitativen Interviewstudie die chemiedidaktischen Elemente einer universitären Lehrerausbildung. Die Entwicklung des Professionswissens wird hierbei als eine aufeinander aufbauende Abfolge des Erwerbs von Reflexionswissen, Entscheidungswissen und Routinen betrachtet (Großebram, 2013; Radtke & Webers, 1998), wobei als Kern der Ausbildung das Reflexionswissen erachtet wird. Dieses besteht einerseits aus deklarativem Wissen über fachdidaktische Inhalte und andererseits aus der Fähigkeit, die Inhalte personenbezogen zu hinterfragen. Ziel war es, für die universitäre Lehrerausbildung das chemiedidaktische Reflexionswissen zu beschreiben. So konnten über 50 Anforderungsbereiche formuliert werden, die sich zukünftig vielfältig nutzen lassen.

Nach Girmes (2006) sind die sechs Kompetenzbereiche Diagnostische Kompetenz, Institutionelle Kompetenz, Curriculare Kompetenz, Methodische Kompetenz, Personell-kulturelle Kompetenz und Reflexive und Evaluative Kompetenz besonders für den ersten Ausbildungsabschnitt von großer Bedeutung. Hierbei fallen unter die zugehörigen Aufgaben des letzten Kompetenzbereichs: „Formen der prozessbegleitenden und ergebnisbezogenen Erhebung von Lern- und Lehrergebnissen zur Reflexion und zum Lernen nutzen und Ergebnisse im Blick auf Kriterien bewerten“ (Girmes, 2006, S. 27).

Ein weiteres Modell, welches explizit das Thema Reflexion inkludiert, ist das hierarchische Strukturmodell der Handlungskompetenz nach Frey und Jung (2011). Sie unterscheiden vier Kompetenzklassen, welche zahlreiche Fähigkeitsdimensionen aufweisen, aus denen sich folglich gewisse Fertigkeiten entwickeln. Reflexion wird dabei der Methodenkompetenz zugeordnet. Diese ist nach Frey, Jäger und Renold (2003) definiert als bestimmte Bereiche, welche eine Person befähigen, innerhalb eines definierten Sachbereichs denk- und handlungsfähig zu sein. Hierunter zählen sie die Fähigkeit, Arbeitsgegenstände zu analysieren, Arbeitsprozesse zu strukturieren über Arbeitsbedingungen und Interaktionspartner sowie über individuell und sozial wirksame Arbeitszusammenhänge zu reflektieren, diese Fähigkeiten einzusetzen, zu erweitern und gegebenenfalls zu modifizieren (Frey, 2004). Methodenkompetenz beinhaltet häufig Fähigkeitsbereiche, die der Analysefähigkeit, Flexibilität, Reflexivität, dem zielorientierten Handeln und den Arbeitstechniken zugeordnet werden (Bader, 2002). Insgesamt sind die vier Kompetenzbereiche keine voneinander inhaltlich und methodisch unabhängigen Kompetenzfacetten, sondern müssen als miteinander vernetzte Kompetenzklassen gesehen werden, die zusammen die Handlungskompetenz einer Person konstruieren, wobei der Aufbau und die Synthese dieser Kompetenzen zum Aufbau einer umfassenden Handlungskompetenz führen (Frey, 2004). Die Fertigkeiten wurden in Abbildung 10 aus Übersichtlichkeitsgründen weggelassen.

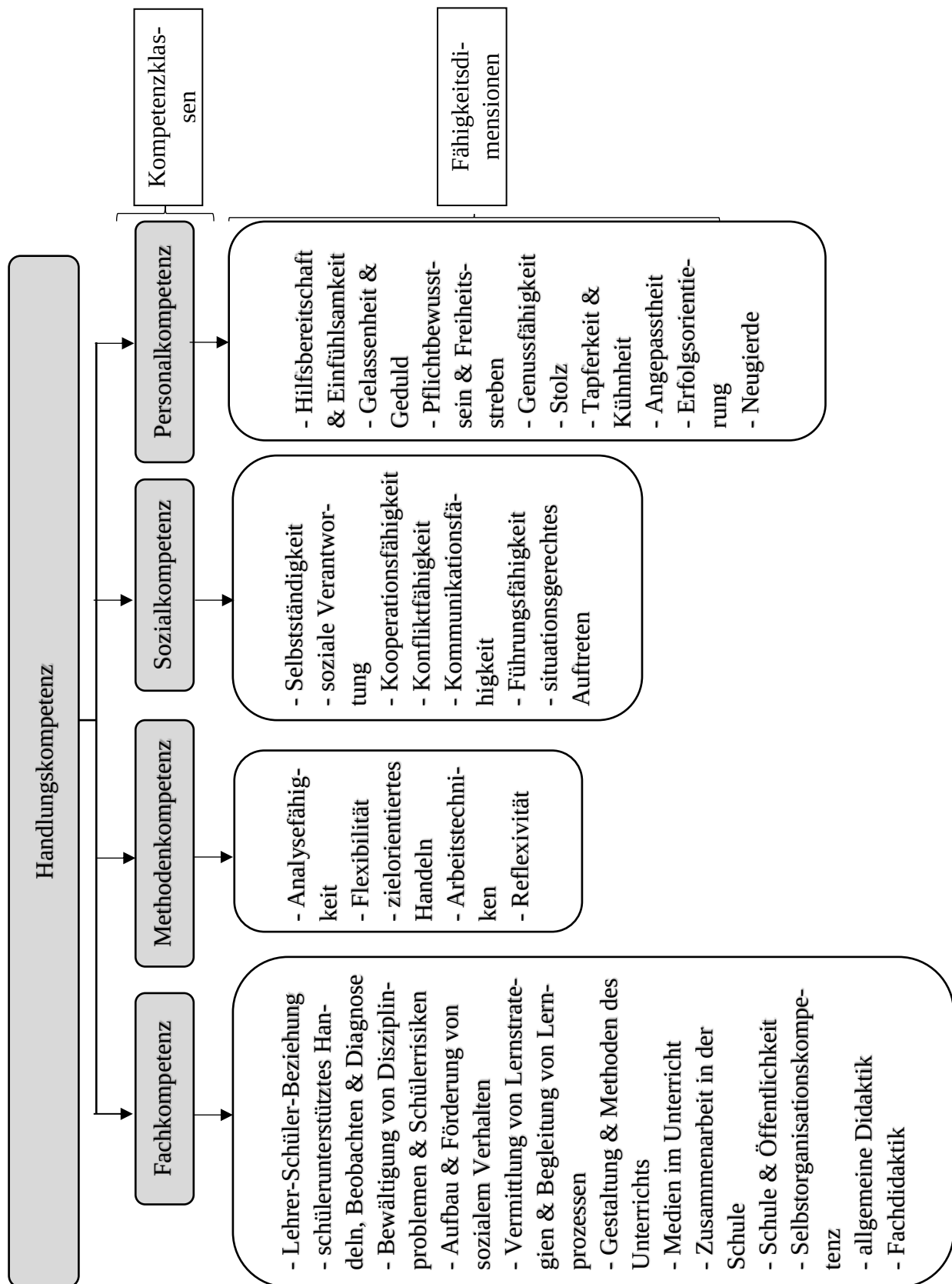


Abbildung 10: Hierarchisches Stufenmodell der Handlungskompetenz (Frey & Jung, 2011, S. 552)

Die zweite Klasse der Kompetenzmodelle bezieht sich auf das Professionswissen: Zunächst ist Lehrerprofessionswissen definiert, als das Wissen, welches Lehrerinnen und Lehrer benötigen, um zu unterrichten bzw. um ihre Profession auszuüben (Schmidt, 2014). Einen interessanten Ansatz der Entwicklung professioneller Handlungskompetenz inklusive Reflexionskompetenz bieten Dreyfus, Dreyfus und Athanasiou (2000) in ihrem Buch „Mind over Machine“. Dieser Ansatz entstammt nicht dem pädagogisch-didaktischen Diskurs, da er sich auf die Grenzen technologischen Denkens bezieht, liefert aber mit dem fünfstufigen Ansatz die Grundlage für das Modell von Koch-Priewe (2002), welche die Stufen ergänzte und so folgendes Experten-Novizen-Paradigma für den Lehrberuf erstellte (Tabelle 4):

Tabelle 4: Experten-Novizen-Paradigma für den Lehrberuf (Koch-Priewe, 2002, S. 317)

Stufen	Merkmale
1. Novize/Novizin	- kontextfreie Regeln werden befolgt, die sich nur auf einzelne Aspekte der Situation beziehen - dies gibt Handlungssicherheit, Verantwortung wird jedoch nicht übernommen unter Berufung auf gelernte Regeln und Prozeduren - das Verhalten ist unflexibel, die einseitige fokussierte Aufmerksamkeit verhindert Multitasking-Fähigkeit
2. Fortgeschrittene(r) Anfänger(in)	- Vermischung von Buchwissen (kontextfreien Regeln) und Berufserfahrung (situationale Regeln) - Aufbau von Fall- oder Episodenwissen - begrifflich differenzierte Wahrnehmung von Situationen - nicht wirklich flexibel
3. Kompetente(r) Lehrer(in)	- keine Orientierung an Handlungsregeln mehr - kann kontextbezogene Ziele nennen, wendet hierarchisch geordnete Entscheidungsprozeduren an - sieht unterschiedliche Wege, Ziele zu erreichen - Flexibel, selbstverantwortlich, bewahrt beim Verstehen und bei Entscheidungen Abstand zu den Dingen, fühlt sich zutiefst verbunden mit den Konsequenzen der eigenen Handlung
4. Gewandt und geübt handelnder(r) Lehrer(in)	- elaborierte Wahrnehmungsmuster, holistisches Verständnis des Prozesses, intuitive Fähigkeit, patterns (Muster) zu benutzen, ohne sie in Komponenten zu zerlegen - erkennt Ähnlichkeit bei scheinbar ganz unterschiedlichen Situationen
5. Experte/Expertin	- holistische Wahrnehmung und integrierte Auffassung von Situationen - reagiert ohne Anstrengung, flüssig, schnell und den jeweiligen situativen Erfordernissen angemessen, durch ein vielseitiges Handlungswissen – Know-how

So entstand ein neues Paradigma, der wissenschaftspsychologische Zugang, dessen Ausarbeitung in Deutschland vor allem durch das Werk von Bromme (2014) gefördert wurde. In

diesem Modell finden, mit Zunahme reflektierter Berufserfahrung, Veränderungsprozesse statt. Lehramtsstudierende befinden sich auf erster Stufe und vollziehen teilweise bis zum Ende des Studiums den Übergang zu Stufe zwei (Denner & Gesenhues, 2013). Indem sie sich meist kontextfrei auf Regeln und Theoriewissen stützen und auf unreflektiert übernommene Regeln und Prozeduren verweisen, erlangen sie zwar Sicherheit, doch dies geht zu Lasten der Verantwortungsübernahme für das eigene Handeln. Denner und Gesenhues (2013) ergänzen zudem, dass Studierende in den ersten Unterrichtsstunden infolge der Konzentration auf den Einsatz von Medien so stark beansprucht werden, dass parallel ablaufende Prozesse nicht mehr wahrgenommen werden können. Dreyfus et al. (2000) merken an, dass Experten intuitiv auf Basis des Erfahrungswissens Entscheidungen treffen, weniger reflexiv. Reflexion beginnt erst, wenn dafür genug Zeit zur Verfügung steht oder das Ergebnis der Handlung oder ihrer Entscheidung kritisch eingestuft wird.

Die zahlreichen verschiedenen Konzeptualisierungen des Professionswissens sind sich weitestgehend einig, dass Lehrkräfte Wissen mit fachspezifischen und didaktisch-pädagogischen Elementen benötigen (Neuweg, 2014). Der kompetenztheoretische Ansatz der Professionsforschung geht auf die Topologie Shulmans (1986) zurück, welcher ein grundlegendes theoretisches Rahmenmodell für viele weitere folgende Modelle entwickelt hat. Dieses beinhaltet die folgenden sieben Bereiche: allgemeines pädagogisches Wissen (general pedagogic knowledge = PK), das Fachwissen (subject matter content knowledge = CK), das fachdidaktische Wissen (pedagogic content knowledge = PCK), das Wissen über das Fachcurriculum (curriculum knowledge), Wissen über die Lernenden und ihre Charakteristika (knowledge of learners and their characteristics“), Wissen über den unterrichtlichen Kontext (knowledge of educational context) und Wissen über die Ziele und Werte von Unterricht (knowledge of educational ends, purposes and values) (Shulman, 1987). Shulman erachtet Selbstreflexion als wesentlichen Bestandteil des Professionalitätsbegriffs: „The teacher is not only a master of procedure but also of content and rationale, and capable of explaining why something is done. The teacher is capable of reflection leading to self-knowledge [...]. A professional is capable not only of practising and understanding his or her craft, but of communicating the reasons for professional decisions and actions to others“ (Shulman, 1986, S. 13). Er erwähnt hier bereits die Fähigkeit eines Lehrers, Entscheidungen begründen zu können und entwickelte neben den sieben Stufen seiner Wissenstypologie ein zyklisches Modell, welches den Reflexionsprozess abbildet. Er schreibt: „[...] pedagogical reasoning and action involve a cycle through the activities of comprehension, transformation, instruction, evaluation and reflection“ (Shulman, 1987, S. 14). Shulman kritisierte, dass in der damaligen Diskussion um den professionellen Lehrer Reflexion schlichtweg ignoriert würde. In seinem Professionsmodell wird dennoch nicht gänzlich klar, wie die sieben Stufen mit Reflexionsfähigkeit in Verbindung zu setzen sind (Roters, 2012). Auch in den National Board for Professional Teaching Standards, also den nationalen Lehrerstandards der USA, diene Shulmans Modell als Basis der fünf Prämissen:

- “Proposition 1: Teachers are committed to students and their learning
- Proposition 2: Teachers know the subjects they teach and how to teach those subjects to students
- Proposition 3: Teachers are responsible for managing and monitoring student learning

- Proposition 4: Teachers think systematically about their practice and learn from experience
- Proposition 5: Teachers are members of learning communities” (Boyd & Reese, 2006, S. 52)

Die Reflexionskompetenz ist dabei im vierten Punkt verankert. 1999 wurden durch Magnusson, Krajcik und Borko die Bereiche von Shulman um die Überzeugungen und Werthaltungen bezüglich des Unterrichtens (beliefs) ergänzt. Als besonders wichtig in den Naturwissenschaften werden die Bereiche Fachwissen, das fachdidaktische Wissen und das pädagogisch-psychologische Wissen erachtet (Abell, Appleton & Hanuscin, 2014; Kunter, Klusmann & Baumert, 2009; Tepner et al., 2012), welche auch als Wissensdomänen in dem Modell professioneller Handlungskompetenz von Baumert und Kunter (2006, S. 482) im Rahmen der COACTIV-Studie verankert sind (Abbildung 11):

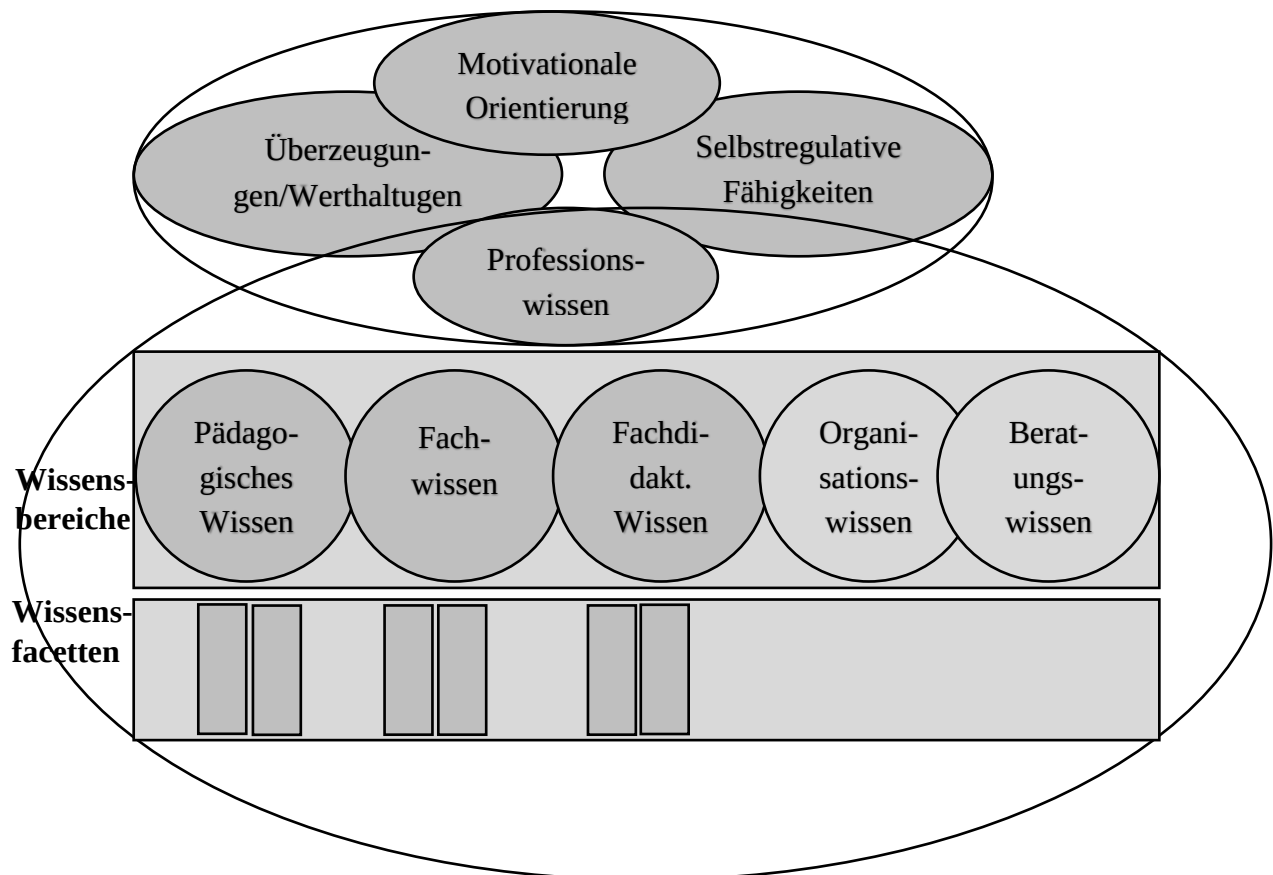


Abbildung 11: Modell professioneller Handlungskompetenz (Baumert & Kunter, 2006, S. 482)

Baumert und Kunter (2006) orientierten sich stark an den fünf Kernaussagen des NBPTS (National Board for Professional Teaching Standards) und übersetzten die vierte Kernaussage („Teachers think systematically about their practice and learn from experience“) als metakognitive Fähigkeiten. Dennoch tauchen diese in obiger Abbildung so nicht wieder auf (Leonhard et al., 2010; Neuweg, 2005). Während die Reflexionskompetenz in diesem Modell nicht explizit erwähnt ist, wird sie von Wyss (2013) den selbstregulativen Fähigkeiten von Lehrpersonen zugeordnet. Sie paraphrasiert nach Baumert und Kunter (2006): „Die Selbstregulationsfähigkeit zeichnet sich aus durch einen verantwortungsvollen Umgang mit den eigenen Ressourcen, sowie einem hohen Berufsengagement gekoppelt mit Distanzierungsfähigkeit“ (Wyss, 2013, S. 60).

Es ist erstaunlich, dass die Reflexionskompetenz bis hierhin nur wenig Einzug in aktuelle Modelle professioneller Handlungskompetenz gefunden hat, obwohl sie in zahlreichen Artikeln und Forschungsarbeiten als unerlässlich für professionelles Lehrerhandeln herausgestellt wird. Herzig, Grafe und Reinhold (S. 150) verstehen professionelles Lehrerhandeln unter anderem als die „Fähigkeit und Bereitschaft, sein eigenes Können in kritischer Distanz zu reflektieren und weiterzuentwickeln“ und charakterisieren die „reflexive [...] Auseinandersetzung mit Handlungsmustern und Interaktionsformen sowie mit didaktischen Vorstellungen [als] eine notwendige Voraussetzung zum Erwerb professionellen Könnens.“ Auch Neuweg (2005, S. 220) definiert Professionalisierung als „Wechselspiel von Einlassung auf Erfahrung, Reflexion und Rückübersetzung in neues Handeln und Erfahren.“

Ausgewählte empirische Ergebnisse zur Reflexion und professionellen Handlungskompetenz:

Motivationale Orientierungen

Nach Baumert und Kunter (2006) umfasst dieser Bereich Konstrukte, welche sich auf die Bereitschaft und Intention, bestimmte Handlungen auszuführen, beziehen. Dabei ist dies immer begründet im Willen, ein bestimmtes Ziel zu erreichen. Dazu zählen die Wissensbereiche Kontrollüberzeugungen, Selbstwirksamkeitserwartung und Lehrerenthusiasmus. Bandura (1997) definiert den Wissensbereich Selbstwirksamkeitserwartung als die subjektive Gewissheit, über die Fähigkeiten und Mittel zu verfügen, um neue und auch schwierige Ziele zu erreichen. Die Zielsetzung und das Niveau des Anspruchs werden dabei durch die Erwartung an die eigene Selbstwirksamkeit bestimmt (Lamprecht, 2011). So werden Anstrengung und Beharrlichkeit sowie die Abschirmung gegen konkurrierende Intentionen und die Verarbeitung von Erfolg und Misserfolg durch beständige Reflexion reguliert (Bandura, 1997; Lamprecht, 2011). Öhlschlegel-Haubrock, Rach und Wolf (2014) fanden heraus, dass bei einem Vergleich der Wirksamkeit eines Mentoringkonzepts mit einem dozentenorientierten Lehrformat in Bezug auf den Wissenszuwachs das klassische Format zwar von den Studierenden besser beurteilt wurde, aber nur beim Mentorenformat eine höhere Einschätzung der Selbstwirksamkeitserwartung und eine verringerte Wahrnehmung der psychischen Belastung beobachtet werden konnte. Sie förderten hierbei das selbstgesteuerte Lernen durch eine adäquate Lernumgebung, welche den Studierenden ermöglichte, selbst aktiv zu werden und kooperativ zu lernen (Öhlschlegel-Haubrock et al., 2014; Reich, 2006), wobei das Lernen unter Peers als besonders effektiv gilt (Nuthall, 2007). Öhlschlegel-Haubrock et al. verwirklichten dies durch studentische Mentorinnen und Mentoren als Betreuer der Kleingruppen. Dohrmann und Nordmeier (2018) untersuchten in ihrer Mixed-Methods-Studie die Auswirkungen des Unterrichtens in Lehr-Lern-Laboren auf Reflexion und Selbstwirksamkeitserwartungen bei Studierenden aus dem Bachelorstudiengang Lehramt Physik und aus dem Bachelorstudiengang Grundschulpädagogik (Sachunterricht). Zu beiden Schwerpunkten wurden Erhebungen im Pre-Post-Design vorgenommen, welche mittels qualitativ gewonnener Interviewdaten erarbeitet wurden. Es konnte eine Milderung des Praxisschocks (nur kleine Effekte im Pre-Post-Vergleich) und zusätzlich eine Steigerung der Selbstwirksamkeitserwartung ($t(69) = 6.14, p < .001$) der Teilnehmer und Teilnehmerinnen erreicht werden. Zusätzlich konnten sie eine positive Einstellung gegenüber dem Reflektieren bestätigen bzw. weiter ausbauen ($t(50) = 3.76; p < .001$). Die Autoren geben jedoch zu bedenken, dass diese Ergebnisse nicht überschätzt werden dürfen, da in einer parallel durchge-

fürten Studie zur Reflexionstiefe (als Teildimension von Reflexionskompetenz) im selben Lehr-Lern-Labor keine signifikanten Änderungen über den Seminarverlauf festgestellt werden konnten (Meißner, Dohrmann & Nordmeier, 2020). Weitere Studien konnten zeigen, dass die Analyse von fremden und eigenen Unterrichtsvideos Lehramtsstudierenden hilft, ihre professionelle Wahrnehmung von Klassenführung (PWKF) und ihre Selbstwirksamkeitserwartungen über Klassenführung (SWKF) zu verbessern (Gold, Hellermann & Holodynski, 2017). Es wurde in dieser Studie außerdem eine positive Korrelation zwischen PWKF und SWKF gefunden. Zudem scheint ein videobasierter Reflexions- und Feedbackzirkel einen motivierenden Faktor und einen positiven Effekt auf die allgemeine Selbstwirksamkeitserwartung der Studierenden zu haben (Weber, Prilop & Kleinknecht, 2019).

Überzeugungen und Werthaltungen

Dieser Bereich bezieht sich nach Baumert und Kunter (2006) auf die Punkte Wertbindung (z. B. Berufsmoral), epistemologische Überzeugungen, subjektive Theorien über Lehren und Lernen und Zielsysteme für Curriculum und Unterricht. Andere subsumieren diese Facette in die Domäne fachdidaktisches Wissen (Borko & Putnam, 1996; Park & Oliver, 2008; Staub & Stern, 2002). Trotzdem haben Überzeugungen zum Lehren von Naturwissenschaften nach Park und Oliver (2008) einen großen Stellenwert für den Unterricht, da sie Entscheidungen hinsichtlich der Auswahl von Instruktionsstrategien und Material sowie hinsichtlich der Bewertung der Schülerleistungen beeinflussen (Borko & Putnam, 1996). Definiert werden Überzeugungen nach Op't Eynde, Corte und Verschaffel als subjektiv für wahr gehaltene Konzeptionen, die die Wahrnehmung und das Handeln beeinflussen. Als starke handlungsleitende Funktionen beschreiben Staub und Stern (2002) diese Domäne. In der Metaanalyse von 93 empirischen Studien zur Entwicklung der Lehrkompetenz fanden Wideen, Mayer-Smith und Moon (1998) wenig Auswirkungen auf epistemologische Überzeugungen von Lehramtsanwärtern bei traditionell ausgerichteten Programmen. Als erfolgreich erwiesen sich jedoch langfristige Zusammenarbeiten zwischen Studierenden und Dozenten bzw. Anwärtern und Ausbildern. Positive Effekte berichteten sie vor allem auch im Zusammenhang mit einer kleinen Klassengröße. Mit niedrigen Teilnehmerzahlen und Lernen in Gruppenarbeiten mit engem Bezug zum Ausbilder waren die Ergebnisse am zufriedenstellendsten. So konnte nachgewiesen werden, dass im Hinblick auf die Überzeugungen der Lehramtsanwärter nicht die Ausbildungslänge oder die Anzahl der Schulpraktika entscheidend war, sondern die Art und Weise der Ausgestaltung der Kurse. Weiter förderlich für kompetenzorientierte Unterrichtsperspektiven waren somit eine kohärente Programmstruktur, Feedback von anderen Studierenden, Mentoren oder Dozenten und strukturierte Praxisphasen. Die Studierenden zeigten sich gegenüber anderen wesentlich experimentierfreudiger in der Unterrichtsplanung:

„In those studies where the candidates were supported by program, peers, and classroom situations, and where deliberative exploration and reflection were encouraged, we saw the flowering of empowered teachers. Here we heard from individuals who experienced the satisfaction of discovering connections between their teaching and their pupils'-learning. These were beginning teachers who were not afraid to experiment, struggle, and make mistakes[...].“ (Wideen et al., 1998, S. 159)

Nach Lamprecht (2011) sind epistemologische Überzeugungen Vorstellungen über das Wissen und den Wissenserwerb einer Fachdisziplin. Aktuelle Lehr-Lernforschung

kommt hierbei zu dem Ergebnis, dass professionelles Handeln beeinflusst wird, indem Lehrpersonen ihre subjektiven Theorien über Lehren und Lernen auf Unterrichtssituationen übertragen, beobachten, interpretieren und reflektieren und so Zielvorgaben und Erwartungen an Schülerinnen und Schüler und an den Unterricht formuliert werden können (Baumert & Kunter, 2006). Auch in der COACTIV-Studie (Brunner et al., 2006) konnte ein Zusammenhang der Überzeugungen mit dem Professionswissen festgestellt werden. Neuhaus (2019) entwickelte ein Seminarkonzept, in dem Studierende zu Beginn des Vorbereitungsseminars ihre Arbeitstheorie sowie die eigenen Erfahrungen formulieren. Im Vorbereitungsjahr werden dann verschiedene Möglichkeiten der Auseinandersetzung mit musikdidaktischen Konzeptionen angeboten. Vor Beginn sowie nach Abschluss des Vorbereitungsseminars formulieren die Studierenden diese Arbeitstheorie erneut und reflektieren deren Veränderungen sowie mögliche Ursachen der Veränderungen. Zur Weiterentwicklung des Formats werden die Arbeitstheorien inhaltsanalytisch ausgewertet. Eine dritte Gruppe startet im Wintersemester 2019/2020 in das Praxissemester; aus diesem Grund können noch keine Ergebnisse berichtet werden. Ziel ist es, die eigenen Überzeugungen zu klären und Veränderungen zu messen.

Selbstregulative Fähigkeiten

Nach Baumert und Kunter (2006) zählen hierzu das persönliche Engagement und die Distanzierungsfähigkeit. Schaarschmidt und Fischer (2003) entwickelten das Testinstrument AVEM (Arbeitsbezogene Verhaltens- und Erlebnismuster), mit welchem sie vier Muster von Lehrerprofilen nachweisen konnten: Muster G (Gesundheit), Muster S (Schutz/Schonung), Risikomuster A (Selbstüberforderung) und Risikomuster B (Resignation/Burn-out). Dabei konnten den Risikomustern A und B als Gemeinsamkeiten innere Unruhe, eingeschränkte Distanzierungsfähigkeit, Unzufriedenheit, Misserfolgserleben und Erleben mangelnder sozialer Unterstützung zugeordnet werden. Das Belastungserleben hat nach einer Studie von Maslach und Leiter (1999) auch Auswirkungen auf das Unterrichtsverhalten und auf die Unterrichtsqualität. Als Maßnahmen der Intervention werden reflexive Überlegungen vorgeschlagen, wie die Klärung des professionellen Selbstverständnisses und die Korrektur unrealistischer und überhöhter (A) bzw. enttäuschter (B) berufsrelevanter Ansprüche, Erwartungen und Zielvorstellungen (Schaarschmidt & Fischer, 2003). Wubbels und Korthagen (1990) zeigten in ihrer Studie, dass Reflexion längerfristige Effekte auf die berufliche Zufriedenheit zur Folge hat. In der Kontrollgruppenuntersuchung mit Lehrkräften konnten sie für die erste Ausbildungsphase keine statistisch bedeutsamen Unterschiede hinsichtlich der Wirksamkeit von Reflexion finden. In den Bereichen Jobzufriedenheit, Angemessenheit der Wahrnehmung unterrichtlichen Handelns und der Beziehungsqualität zu den Schülerinnen und Schülern ließen sich jedoch positive Ergebnisse erzielen. Ergänzend konnten Klusmann, Kunter, Trautwein und Baumert (2006) zeigen, dass hohe selbstregulative Fähigkeiten mit großem Engagement der Lehrkraft aber auch mit der Fähigkeit verbunden ist, den notwendigen Abstand zum Berufsalltag aufbauen zu können. Brouër und Gläser-Zikuda (2010) berichten aus ihren Hochschulprojekten zur Förderung selbstregulativer Fähigkeiten von Studierenden ernüchternde Ergebnisse. Im Zentrum steht hier die Arbeit mit Lerntagebüchern und Portfolios, welche in der Literatur als förderlich für selbstregulative Fähigkeiten gelten. Jedoch wird in beiden Studien die Selbstreflexion der Studierenden als lediglich oberflächlich berichtet. Zudem wird das Verfassen von Portfolios bzw. Lerntagebüchern als arbeitsaufwendig eingestuft und tendenziell negativ bewertet.

Professionswissen

Nach Baumert und Kunter (2006) lässt sich Professionswissen in die Wissensbereiche Fachwissen, Fachdidaktisches Wissen, Pädagogisches Wissen, Organisationswissen und Beratungswissen unterteilen. Schwerpunkte bilden jedoch die Betrachtung der drei Wissensbereiche Pädagogisches Wissen, Fachwissen und Fachdidaktisches Wissen, welche die bislang allgemein akzeptierten Kernkategorien des Professionswissens von Lehrkräften darstellen (Baumert & Kunter, 2006; Blömeke, Kaiser & Lehmann, 2008; Krauss et al., 2008; Lipowsky, 2006; Terhart, Bennewitz & Rothland, 2014). Der Zusammenhang zur Reflexionskompetenz und zu anderen Facetten dieser Studie wird im Folgenden genauer dargestellt. Hierbei kann Reflexion als Konstrukt nur über Rückschlüsse auf mentale Prozesse erforscht werden (Hatton & Smith, 1995; Korthagen et al., 2002a). Als beste Methode zur Messung der Reflexionskompetenz gelten schriftliche Berichte über Unterricht und Aufzeichnungen von Dialogen (Danielson, 2008).

Pädagogisches Wissen

Dieser Wissensbereich beinhaltet allgemein pädagogische Inhalte und ist somit für alle Unterrichtsfächer gültig, also fachindifferent (Grossman, 1990; König & Blömeke, 2009; König & Seifert, 2012a; Neuweg, 2011; Voss & Kunter, 2011). Jedoch herrscht Uneinigkeit darüber, aus welchen Facetten pädagogisches Wissen genau besteht (Riese, 2009). In Tabelle 5 wird eine Übersicht über ausgewählte Definitionen des pädagogischen Wissens gegeben:

Tabelle 5: Ausgewählte Definitionen zum Konstrukt „pädagogisches Wissens“

Name	Studie	Definition
Shulman (1987)		„general pedagogical knowledge, with special reference to those broad principles and strategies of classroom management and organization that appear to transcend subject matter“ (Shulman, 1987, S. 8)
Borko und Putnam (1996)		„a teacher’s knowledge and beliefs about teaching, learning, and learners that transcend particular subject matter domain. [...] knowledge of various strategies and arrangements for effective classroom management, instructional strategies for conducting lessons and creating learning environments, and more fundamental knowledge and beliefs about learners, how they learn, and how that learning can be fostered by teaching“ (Borko & Putnam, 1996, S. 675)
Blömeke, Felbrich und Müller (2008)	MT21 (Mathematics Teaching in the 21st Century)	- Umgang mit sozialer Ungerechtigkeit - Lernzielkontrolle - Unterrichtsplanung

(Blömeke, Kaiser, Lehmann et al., 2008)	TEDS-M (Teacher Education and Development: Learning to teach Mathematics)	- Umgang mit Heterogenität - Leistungsbeurteilung - Strukturierung von Unterricht - Klassenführung - Motivierung
Seifert, Hilligus und Schaper (2009)	SPEE (Standards – Profile – Entwicklung – Evaluation)	Inhaltsbereiche: - Erziehung und Bildung - Unterricht und allgemeine Didaktik - Schulentwicklung und Gesellschaft Kognitive Anforderungsbereiche: - Wissen reproduzieren und verstehen - reflektieren und anwenden, urteilen - bewerten und entscheiden
Kunter, Baumert, Blum et al. (2011)	COACTIV (Cognitive Activation in the Classroom)	- Wissen über Lernprozesse - Wissen um Leistungsbeurteilung - Wissen über effektive Klassenführung
Voss und Kunter (2011)	COACTIV-R	- Wissen über Schülerinnen und Schüler - Wissen über Leistungsbeurteilung - Wissen über effektive Klassenführung - Wissen über Unterrichtsmethoden
Borowski et al. (2010)	ProwiN (Professionswissen in den Naturwissenschaften)	- Klassenführung - Unterrichtsmethoden - individuelle Lernprozesse - Leistungsbeurteilung
König und Seifert (2012a)	LEK (Längsschnittliche Erhebung pädagogischer Kompetenzen von Lehramtsstudierenden)	- Umgang mit Heterogenität - Leistungsbeurteilung - Strukturierung von Unterricht - Klassenführung/Motivation - Unterricht und allgemeine Didaktik - Erziehung und Bildung - Schulentwicklung
Kunina-Habenicht et al. (2012)	BilWiss (Bildungswissenschaftliches Wissen und der Erwerb professioneller Kompetenz in der Lehramtsausbildung)	- Heterogenität und soziale Funktion - Diagnostik und Evaluation - Unterricht - Bildungstheorie - Bildungssystem und Schulorganisation - Entwicklungsprozesse - Lehrerberuf - Lernprozesse - Sozialisationsprozesse

Borko und Putnam (1996) konnten zeigen, dass pädagogisches Wissen hauptsächlich im Studium erworben und durch praktische Erfahrung automatisiert und erweitert wird. Das Projekt SPEE erzielte in der Pilotierung gute interne Konsistenzen der Gesamtskala, allerdings wurden keine Analysen zur Dimensionalität und internen Konsistenzen der Subskalen berichtet. In einer bundesweiten Studie wurde der Test bei rund 430 Lehramtsstu-

dierenden der Physik aller Semester eingesetzt und konnte ebenfalls gute interne Konsistenzen der Gesamtskala erzielen. Zusätzlich wiesen Riese und Reinhold (2010) nach, dass das pädagogisch-psychologische Wissen das fachdidaktische Wissen der Physiklehramtsstudierenden positiv beeinflusst. Die LEK-Studie (bestehend u. a. aus dem SPEE-Test) konnte zeigen, dass das pädagogische Professionswissen der Lehramtsstudierenden im Verlauf von vier Semestern signifikant zunimmt (die Effektstärken d liegen im Bereich zwischen 0,47 und 1,22; König & Seifert, 2012a, 2012b; Seifert & Schaper, 2012). Dies gilt besonders für den unterrichtsbezogenen Kontext. Ebenso existieren Studien, welche Auswirkungen auf die Schülerleistung belegen (Darling-Hammond, Berry & Thoreson, 2001; Lipowsky, 2006). Es wird bei allen beschriebenen Modellen angenommen, dass das dargestellte Wissen eine Voraussetzung für erfolgreiches Unterrichten ist (Vogelsang, 2014). Lüders (2012) kritisiert an den bestehenden Modellen die problematische Validierung, begründet durch eine unscharfe theoretische Herleitung. Weitere Studien zur Entwicklung der Unterrichtswahrnehmung, welche eng mit der Klassenführung verbunden ist, zeigen eine Zunahme pädagogischen Wissens über Klassenführung und eine Verbesserung der professionellen Unterrichtswahrnehmung durch Trainings – sowohl während des Studiums (Stürmer, Könings & Seidel, 2013) als auch beim Übergang vom Studium zum Beruf (König, 2014). Eine ausführliche Zusammenfassung der Studien zum pädagogischen Wissen und dessen Auswirkungen auf die unterschiedlichsten Bereiche findet sich bei Voss, Kunina-Habenicht, Hoehne und Kunter (2015).

Fachwissen

Um Schülerinnen und Schülern Fachinhalte so näherbringen zu können, dass diese sie verstehen, ist ausreichend Fachwissen als Kompetenz einer Lehrkraft unumgänglich (Kirschner, 2013; Loucks-Horsley & Matsumoto, 1999; Riese, 2009). So ist es das Wissen über Inhalte, Zusammenhänge und Methoden des Faches, dessen konzeptionelles Verständnis bei Lehrkräften vorhanden sein muss, um den Unterricht erfolgreich vorbereiten und durchführen zu können (Blömeke, Kaiser & Lehmann, 2008; Byers, Koba, Sherman, Schepke & Bolus, 2011; Neuweg, 2011; Shulman, 1987). Baxter und Lederman (1999, S. 148) definieren Fachwissen als „the knowledge held by a content expert, what the research chemist understand about the discipline of chemistry.“

In Tabelle 6 wird eine Übersicht über ausgewählte Konzeptualisierungen des Fachwissens gegeben:

Tabelle 6: Ausgewählte Konzeptualisierungen des Fachwissens

Studie/Name	Facetten			
Ball, Hill und Bass (2005); Ball, Lubienski und Mewborn (2001)	Allgemein fachliches Wissen, über das gut ausgebildete Erwachsene verfügen sollten	Wissen, welches speziell für die Ausübung des Lehrberufs erforderlich ist		
Riese und Reinhold (2008)	Wissen auf Schulniveau	Vertieftes Wissen	Universitäres Wissen	

MT21-Studie (Blömeke, Kaiser & Lehmann, 2008)	Mathematik der Sekundarstufe I	Mathematik der Sekundarstufe II	Schulmathe- matik vom höheren Stand- punkt aus	Universitäre Mathematik
TEDS-M (Blömeke, Kaiser, Leh- mann et al., 2008)	Elementares Niveau (Pri- marstufe)	Mittleres Niveau (Sekundarstufe I)	Fortgeschritte- nes Niveau (höhere Klas- senstufe)	
COACTIV (Baumert & Kunter, 2006)	Mathe- matisches Alltagswissen	Schulwissen	Vertieftes schulisches Wissen	Akad- emisches Forschung- swissen
ProwiN (Tepner et al., 2012)	Sekundarstufe I-Niveau	Sekundarstufe II- Niveau	Grundstudi- umsniveau	

Zusammenfassend lässt sich festhalten, dass für eine erfolgreiche Vorbereitung auf den Unterricht Wissen erforderlich ist, welches über das zu unterrichtende Schulniveau hinausgeht (Blömeke, Kaiser, Lehmann et al., 2008; Tepner et al., 2012). Auf Grundlage des dreidimensionalen Modells von Tepner et al. (2012, S. 19) wurden im ProwiN Projekt strukturell vergleichbare Tests zum Fachwissen sowie zum pädagogischen und fachdidaktischen Wissen für die Fächer Biologie, Chemie und Physik erarbeitet (Abbildung 12).

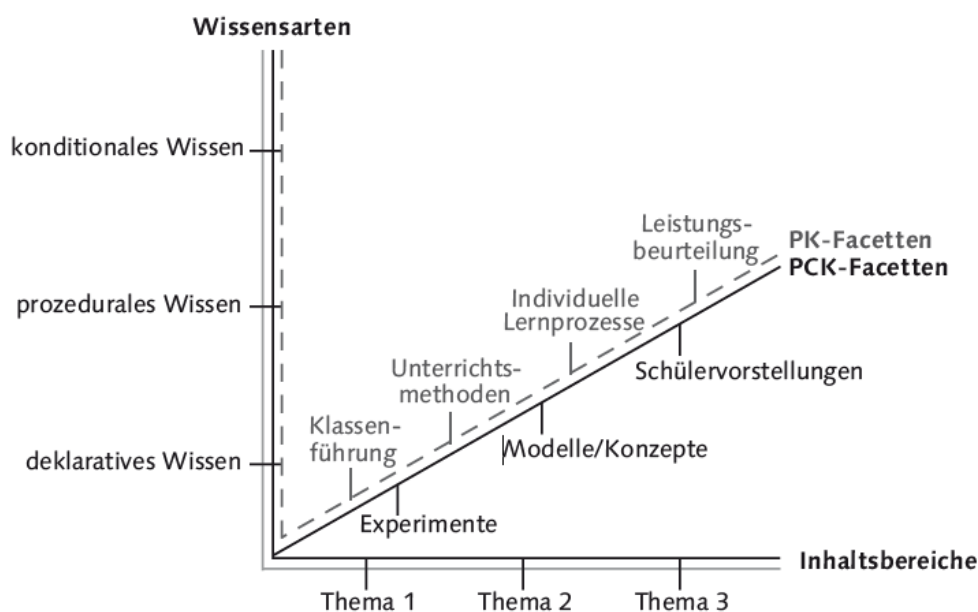


Abbildung 12: Modell zur Konzeption von Items zur Erfassung des Professionswissens (Tepner et al., 2012, S. 19)

Die domänenspezifischen Dimensionen Fachwissen und fachdidaktisches Wissen lassen sich wie bei Blömeke, Kaiser und Lehmann (2008) strukturell analog konzeptualisieren, sodass ein relativer Vergleich des Wissens einer Lehrkraft möglich ist. Es wird, entsprechend dem fachdidaktischen Wissen, das Fachwissen in unterschiedliche Inhaltsbereiche und in die drei Wissensarten deklaratives Wissen, prozedurales Wissen und konditionales

Wissen unterteilt. Letzteres wurde ergänzt, um dem Aspekt der situationsabhängigen Entscheidung bzw. des Urteilens in einer Unterrichtssituation gerecht zu werden. Dies ist bedingt vergleichbar mit Donald Schöns reflection-in-action. Die Autoren verstehen darunter das Wissen über Bedingungen, unter denen eine Entscheidung bzw. eine Handlung angemessen ist (Tepner et al., 2012, S. 17). Zudem umfasst es das Wissen, welches benötigt wird, um Prozesse und Handlungen planen und begründen zu können. So können die drei Bereiche wie folgt zusammengefasst werden:

- Konditionales Wissen: Wissen, wann und warum ein Prozess oder eine Handlung angewandt werden kann
- Prozedurales Wissen: Wissen, wie etwas abläuft
- Deklaratives Wissen: Kenntnis, dass bzw. was etwas ist; Wissen über Begriffe, Tatsachen und Prinzipien (Anderson, 1988; Gruber, 2008; König, 2014; Paris, Lipson & Wixson, 1983)

Fachdidaktisches Wissen

Nach Shulman (1986) besitzen Lehrerinnen und Lehrer spezifisches Wissen, um Fachinhalte adressatengerecht an Schülerinnen und Schüler zu vermitteln, das sogenannte pedagogical content knowledge. Es ist ein Zusammenspiel von pädagogischem Wissen und Fachwissen und wird im deutschsprachigen Raum als fachdidaktisches Wissen bezeichnet (Baumert & Kunter, 2006; Blömeke, Kaiser & Lehmann, 2008; Lipowsky, 2006). In Anlehnung an COACTIV beinhaltet das fachdidaktische Wissen die beide Facetten Wissen über Schülerkognitionen (Schülerfehler/Schülervorstellungen) und Wissen über instruktionale Strategien (Abell, Rogers, Hanuscin, Lee & Gagnon, 2009; Shulman, 1986). In jüngster Forschung forcierten die Arbeiten mit dem Thema professionelle Handlungskompetenz von Lehrkräften den Schwerpunkt fachdidaktisches Wissen (PCK). So entstand ein Modell für das fachdidaktische Wissen von Lehrkräften in der Zusammenarbeit von 24 internationalen Forschern, das sogenannte Refined Consensus Model (RCM). Es wurde zum einen angelehnt an das Modell professioneller Handlungskompetenz von 2012 (das 2012 Consensus Model (CM)), welches PCK inkludierte. Zum anderen wurde die Definition von PCK als „the knowledge of, reasoning behind, and planning for teaching a particular topic in a particular way for a particular purpose to particular students for enhanced student outcomes (Reflection *on* Action, Explicit)“ (Gess-Newsome, 2015, S. 45) Grundlage der Überarbeitung. Gess-Newsome (2015) unterscheidet weiter: „Personal PCK&S is the act of teaching a particular topic in a particular way for a particular purpose to particular students for enhanced student outcomes (Reflection *in* Action, Tacit or Explicit).“ Auch hier wird erneut der Bezug zu Donald Schöns Unterscheidung der reflection on action und der reflection in action deutlich. Carlson und Daehler (2019) bezeichnen ihr Modell als alles andere als perfekt, aber nichtsdestotrotz als beste Repräsentation und Beschreibung des bisherigen Standes der Forschung. Dabei wird PCK in drei Bereiche unterteilt: das kollektive PCK (collective PCK (cPCK)), das persönliche PCK (personal PCK (pPCK)) und das verordnete PCK (enacted PCK (ePCK)). Ersteres beschreibt hier das spezialisierte, professionelle Wissen, welches öffentlich und kollektiv ist, also beispielsweise in wissenschaftlichen Kreisen dokumentiert ist und z. B. in universitären Lehrveranstaltungen kommuniziert wird. Zweiteres ist das persönliche Expertenwissen einer einzelnen Lehrkraft in den Naturwissenschaften und Letzteres verkörpert dasjenige Wissen, welches zu pädagogischen Begründungen und Schlussfolgerungen während der Planungsphasen, des Unterrichtens und der anschließenden Reflexion einer

Stunde führt. Tabelle 7 gibt einen Überblick über die wichtigsten Definitionen des Modells:

Tabelle 7: Wichtige Definitionen des RCM nach Carlson & Daehler (2019)..

cPCK	„The knowledge held by a group of people and considered generalisable to some degree, which is why this layer is situated after the learning context layer.“ (Carlson & Daehler, 2019, S. 89)
pPCK	„A teacher’s personal knowledge and unique expertise about teaching a given subject area, resulting from the cumulative experiences with and contributions from students, peers, and others.“ (Carlson & Daehler, 2019, S. 86)
ePCK	„Specific knowledge and skills utilised by an individual teacher in a particular setting, with a particular student or group of students, with a goal for those students to learn a particular concept, collection of concepts, or a particular aspect of the discipline.“ (Carlson & Daehler, 2019, S. 83 f.)
Learning Context	„A multitude of factors that define and mediate learning, including everything from the broader educational climate to individual student attributes.“ (Carlson & Daehler, 2019, S. 87)

In Tabelle 8 wird eine Zusammenfassung ausgewählter Konzeptualisierungen des fachdidaktischen Wissens vorgestellt.

Tabelle 8: Zusammenfassung ausgewählter Konzeptualisierungen des fachdidaktischen Wissens

Studie/ Name	Facetten				
Shulman; Shulman (1986, 1987)	Conceptions and preconceptions of students	Representations, analogies, illustrations, examples, explanations and demonstrations			
Mag-nusson et al. (1999)	Knowledge and beliefs about students’ understanding of specific science topics	Knowledge and beliefs about instructional strategies for teaching science	Knowledge and beliefs about science curriculum	Orienta-tion to-ward teaching science	Knowledge and beliefs about assessment in science
COAC-TIV Baumert und Kunter (2006)	Wissen über Schülervorstellungen und Schülerkognition	Wissen über Erklärungen und Repräsentationsformen	Wissen über das Potential von Aufgaben		

Riese und Reinhold (2010)	Wissen über (allgemeine) Aspekte physikalischer Lernprozesse	Gestaltung und Planung von Lernprozessen	Beurteilung, Analyse und Reflexion von Lernprozessen	Wissen über den Einsatz von Experimenten	Adäquate Reaktion in kritischen, unerwarteten Unterrichtssituationen
ProwiN Tepner et al. (2012)	Schülervorstellungen	Modelle	Experimente		
RCM (Carlson & Daehler, 2019)	Collective PCK	Personal PCK	Enacted PCK Planung, Unterricht, Reflexion		

Alle Modelle weisen zwei Gemeinsamkeiten auf: Die kognitiven Strukturen von Schülerinnen und Schülern, welche das Wissen über Schülervorstellungen bzw. Bedingungen des Lernens sowie die Interaktion mit Schülerinnen und Schülern inkludieren sowie das Wissen über Elemente des Unterrichtens wie Erklärungen, Repräsentationsformen und Instruktionsstrategien (Baumert & Kunter, 2006; Hill, Ball & Schilling, 2008; Schmidt, 2014; Tepner et al., 2012). Anthofer (2017) gibt jedoch zu bedenken, dass Übersichten in ihrer Aussagekraft begrenzt sind, da die Facettenbezeichnungen mit unterschiedlichen Konstrukten verbunden sein können. Ein vollständiger Vergleich ist nur mit intensiver Auseinandersetzung der Rahmenkonzeptionen möglich.

So wird im Refined Consensus Model (RCM) deutlich (Abbildung 13), dass die Basis professionellen Wissens (gemeint sind Fachwissen und Pädagogisches Wissen) fundamental für das fachdidaktische Wissen einer Lehrkraft der Naturwissenschaften ist (Borowski, Kirschner, Liedtke & Fischer, 2011; Kunter, Baumert & Blum, 2011; Magnusson et al., 1999) und der Lernkontext (gemeint sind Schulpolitik, Wertegemeinschaften und die Merkmalsausprägungen der Schülerinnen und Schüler) das Lehren und Lernen beeinflussen. Durch Wechselwirkungen der im Modell als konzentrische Kreise dargestellten Ebenen wird das persönliche PCK beeinflusst. Die Auswirkung des Fachwissens auf das PCK wurde bereits nachgewiesen (Käpylä, Heikkinen & Asunta, 2009; van Driel, Verloop & Vos, 1998). So lassen sich Selbstwirksamkeitserwartungen, Zielorientierungen und fachspezifische Überzeugungen, als Beispiele für Überzeugungen, vertiefen und/oder filtern bzw. reflektieren und so das pPCK weiterentwickeln. Ebenso kann pPCK durch professionelles Feedback gefördert werden.

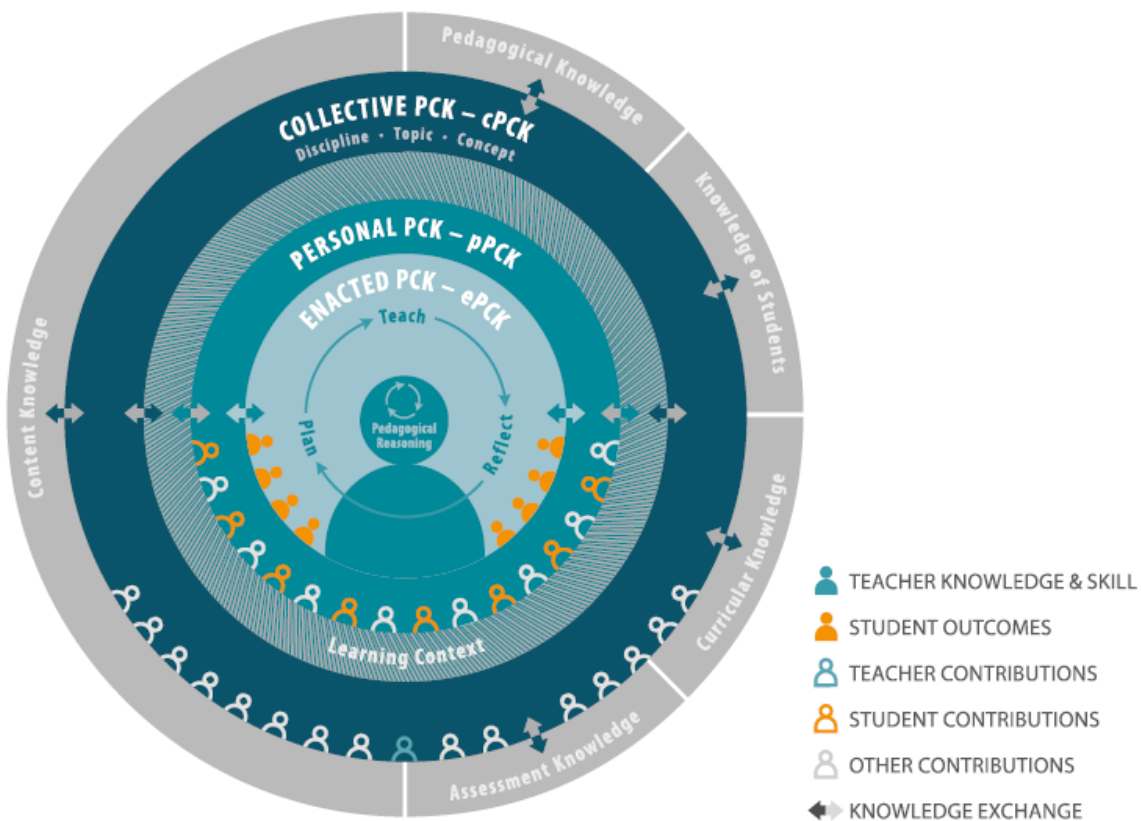


Abbildung 13: Das Refined Consensus Model (RCM) of PCK (Carlson & Daehler, 2019, S. 83)

Um den Bezug zur Reflexionskompetenz zu schaffen, wird im Folgenden näher auf das enacted PCK (ePCK) eingegangen (Abbildung 14):

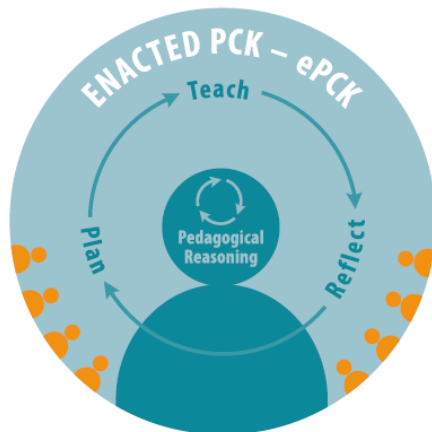


Abbildung 14: Enacted PCK (Carlson & Daehler, 2019, S. 84)

ePCK wird definiert als „specific knowledge and skills utilised by an individual teacher in a particular setting, with a particular student or group of students, with a goal for those students to learn a particular concept, collection of concepts, or a particular aspect of the discipline“ (Carlson & Daehler, 2019, S. 83 f.). Dabei nehmen die Autoren Bezug auf Schön (1984) und seine Unterscheidung der reflection in action und reflection on action und betonen, dass die Anwendung des Modells nicht nur von dem Wissen über Lehrerverhalten und Begründungen der Lehrerhandlungen in der Interaktion mit den Schülerinnen und Schülern abhängt, sondern ebenso von der Unterrichtsplanung und der Reflexion über Planung und Wirksamkeit.

Das ePCK einer Lehrperson reflektiert zu jedem beliebigen Unterrichtszeitpunkt den Kontext der Schule und des Klassenzimmers und jeder einzelne Schüler, jede einzelne Schülerin interagiert mit der Lehrperson und ihrem Verständnis des naturwissenschaftlichen Sachverhalts sowie ihren pädagogischen Kenntnissen und Fähigkeiten. Eine Schlüsselrolle spielt hier die Art und Weise wie eine Lehrkraft auf ihr Wissen zurückgreift, um die individuellen Bedürfnisse der Schülerinnen und Schüler

während des Unterrichts zu erkennen. Eine Lehrperson muss die Zeit haben zu planen, zu lehren und zu reflektieren. Dies geschieht idealerweise in Kooperation mit anderen Lehrkräften. Zusammengefasst bedeutet das, dass ePCK sichtbar wird durch den Wissensstand einer Lehrkraft, durch die Wahl der Unterrichtsmethoden und -strategien, in der Begründung für spezifische pädagogische Handlungen und in der Integration multipler Faktoren pädagogischen Handelns (z. B. Kenntnisse der Schülerinnen und Schüler oder Bewertungskompetenz). Zur genaueren Ausarbeitung des ePCK schlagen Carlson und Daehler (2019) vor, zu untersuchen, wie Lehrpersonen ihren Unterricht bezüglich eines bestimmten Themas planen, reflektieren und unterrichten. Dies geschieht in dieser Studie.

Schmelzing, Wüsten, Sandmann und Neuhaus (2010) unterteilen das fachdidaktische Wissen in deklaratives, prozedurales und reflexives Wissen. Dabei wird reflexives Wissen, vergleichbar mit Donald Schöns reflection on action, nach dem Handeln mittels Videovignetten abgerufen. Reflexives Wissen bezieht sich hier auf bestimmte Situationen und Bedingungen und betrifft das „Wann“ und „Warum“ einer Handlung. Unter anderem verfolgte Schmelzing das Ziel, Skalen zur Erfassung reflexiver Komponenten des fachdidaktischen Wissens zu entwickeln, evaluieren und validieren. Die reflexiven Anteile wurden mithilfe eines videogestützten Paper-Pencil-Tests erfasst, bestehend aus authentischen Unterrichtssituationen in Form von Videovignetten, zu welchen die Probandinnen und Probanden Stellung beziehen sollten. Folgende Facetten wurden von dem Test erfasst: der Aufmerksamkeitsfokus, die Bewertung der dargestellten Unterrichtssituation, fachdidaktische Orientierungen und die Fertigkeit, fachdidaktisch und im Rückgriff auf entsprechende Theorien und Konzepte zu argumentieren. Die Validität des Messinstruments ließ sich für Biologie-Lehramtsstudierende sowie Referendarinnen und Referendare nachweisen (für erfahrene Biologielehrkräfte konnte diese jedoch nicht für alle Subskalen eindeutig geklärt werden).

Fazit:

Es ist festzuhalten, dass sich Lehrerprofessionalität durch Routinen und Fertigkeiten basierend auf Erfahrungswissen einerseits, sowie dessen Reflexion und Analyse andererseits auszeichnet (Combe & Kolb, 2004). Aktuelle Forschungsarbeiten betonen das Prozesshafte einer beständigen Weiterentwicklung und eine erforderliche Einsicht in die Notwendigkeit zu lebenslangem Lernen (Helsper & Tippelt, 2011; Mönig, 2013; Nittel, 2011). Als „zentrale Vermittlungsinstanz zwischen Wissen, Können und Erfahrung“ (Artmann, Herzmann, Hoffmann & Proske, 2013, 134) ist die Fähigkeit des Reflektierens eine gedankliche Strukturierung der Erfahrungen und der wissenschaftlichen Theorien, welche kritisch gespiegelt werden und neues Handeln hervorruft (Leonhard, 2008). Es lassen sich folgende Befunde zum Reflektieren in Unterrichtsprozessen zusammenfassen:

- Studierende erlangen Sicherheit, indem sie sich meist kontextfrei auf Regeln und Theoriewissen stützen und auf unreflektiert übernommene Regeln und Prozeduren verweisen (Denner & Gesenhues, 2013).
- Dies geht zu Lasten der Verantwortungsübernahme für das eigene Handeln (ebd.).
- Studierende in den ersten Unterrichtsstunden sind infolge der Konzentration auf den Einsatz der Medien so stark beansprucht, dass parallel ablaufende Prozesse nicht mehr wahrgenommen werden können (ebd.).
- Reflexion beginnt erst, wenn Zeit vorhanden ist oder wenn das Ergebnis der Handlung oder ihrer Entscheidung kritisch eingestuft wird (Dreyfus et al., 2000).

- Eine professionelle Lehrperson ist fähig, Entscheidungen begründen zu können (Shulman, 1986).
- Ein Mentorenformat hat positive Auswirkungen auf die Selbstwirksamkeitserwartung und eine verringerte Wahrnehmung der psychischen Belastung zur Folge (Öhlschlegel-Haubrock et al., 2014).
- Das Unterrichten in Lehr-Lern-Laboren mit anschließender Reflexion führt zu einer Steigerung der Selbstwirksamkeitserwartung (Dohrmann & Nordmeier, 2018).
- Die Analyse von fremden und eigenen Unterrichtsvideos hilft Lehramtsstudierenden, ihre professionelle Wahrnehmung von Klassenführung und ihre Selbstwirksamkeitserwartungen über Klassenführung zu verbessern (Gold et al., 2017).
- Ein videobasierter Reflexions- und Feedbackzirkel hat einen motivierenden Charakter und einen positiven Effekt auf die allgemeine Selbstwirksamkeitserwartung der Studierenden (Weber et al., 2019).
- Förderlich für kompetenzorientierte Unterrichtsperspektiven sind eine kohärente Programmstruktur, Feedback von anderen Studierenden, Mentoren oder Dozenten und strukturierte Praxisphasen (Wideen et al., 1998).
- Portfolios bzw. Lerntagebücher zur Förderung selbstregulativer Fähigkeiten von Studierenden sind wenig zielführend. Zudem wird das Verfassen von Portfolios bzw. Lerntagebüchern als arbeitsaufwendig eingestuft und tendenziell negativ bewertet (Brouër & Gläser-Zikuda, 2010).
- Als ebenso geeignete Methode zur Messung der Reflexionskompetenz gelten schriftliche Berichte über Unterricht und Aufzeichnungen von Dialogen (Danielson, 2008).
- Für eine erfolgreiche Vorbereitung auf den Unterricht ist Fachwissen erforderlich, welches über das zu unterrichtende Schulniveau hinausgeht (Blömeke, Kaiser, Lehmann et al., 2008; Tepner et al., 2012).
- Enacted PCK verkörpert nach Gess-Newsome et al. (2017) das Wissen, welches zu pädagogischen Begründungen und Schlussfolgerungen während der Planungsphasen, des Unterrichtens und der anschließenden Reflexion einer Stunde führt.
- Die Basis professionellen Wissens (gemeint sind Fachwissen und Pädagogisches Wissen) sind fundamental für das fachdidaktische Wissen einer Lehrkraft der Naturwissenschaften (auch Borowski et al., 2011; Kunter, Baumert & Blum, 2011; Magnusson et al., 1999).
- Das pPCK wird weiterentwickelt, indem Selbstwirksamkeitserwartungen, Zielorientierungen und fachspezifische Überzeugungen vertieft und/oder gefiltert bzw. reflektiert werden (Carlson & Daehler, 2019).
- pPCK kann durch professionelles Feedback gefördert werden (Carlson & Daehler, 2019).
- Eine Lehrperson muss die Zeit haben zu planen, zu lehren und zu reflektieren. Idealerweise verläuft dieser Aspekt bei der Zusammenarbeit mit anderen (Carlson & Daehler, 2019).
- Das ePCK wird sichtbar durch den Wissensstand der Lehrperson, durch die Wahl der Unterrichtsmethoden und -strategien, in der Begründung für spezifische pädagogische Handlungen und in der Integration multipler Faktoren pädagogischen Handelns (z. B. Kenntnisse der Schülerinnen und Schüler oder Bewertungskompetenz) (Carlson & Daehler, 2019).

- Zur genaueren Ausarbeitung des ePCK schlagen Carlson und Daehler (2019) vor, zu untersuchen, wie Lehrpersonen ihren Unterricht bezogen auf ein bestimmtes Thema planen, reflektieren und unterrichten.

3.2.2. Reflexionsfähigkeit – Voraussetzungen

Nach Schädlich (2015, S. 258) zeigt sich reflexive Handlungskompetenz darin, „dass Studierende in der Lage sind, vor dem Hintergrund fachdidaktischer und curricularer Texte (Theorien) zu planen und durchzuführen sowie dabei die Relevanz dieser Texte für die individuellen und komplexen Erfahrungen in der Handlungssituation (Praxis) explizierend diskutieren zu können. Erkennbar wird reflexive Handlungskompetenz in der Performanz rückblickend versprachlichter Handlungslogiken.“ Viele Studien belegen, dass Lehrpersonen sehr wenig reflektieren (El-Dib, 2007; Haas, 1998; Lee, 2005; Marcos, Sanchez & Tillema, 2011). Weiterhin auffällig ist, dass in vielen Studien zum Thema Reflexionskompetenz kaum das höchste Niveau erreicht wird (Abels, 2011; Gore & Zeichner, 1991; Meyer, 2003). Gründe hierfür können in den Professionalisierungsprogrammen gesucht werden (Abels, 2011), in Schwächen in den Erhebungs- und Auswertungsmethoden (Sparks-Langer, Simmons, Pasch, Colton & Starko, 1990) oder in fehlenden Voraussetzungen für wissenschaftliche Reflexionsfähigkeit. Welche Facetten genau diese Kompetenz beinhaltet, soll im Folgenden näher beleuchtet werden.

Nach Dewey umfassen die Voraussetzungen für reflexives Denken drei Hauptkomponenten (Dewey, 1933). Die erste beinhaltet den Prozess, sprich eine systematische Art und Weise des Denkens und ein Endprodukt, oder wie er es bezeichnet, ein Stellungbeziehen. Die zweite Komponente beinhaltet eine spezifische Reflexionsbereitschaft bestehend aus drei Bereichen: Aufgeschlossenheit (open-mindedness), Herzlichkeit (whole-heartedness) und Verantwortung (responsibility). Diese Bereiche sollen interpretiert werden als aktives Verlangen, bestimmte Probleme aus verschiedene Perspektiven zu betrachten (Farrell, 2014). So fordert Dewey: „Knowledge of the methods alone will not suffice; there must be the desire, the will to employ them“ (Dewey, 1933, S. 30). So beschreibt er den Begriff „open-mindedness“ als Bereitschaft, sich mit neuen Ideen und Erklärungsansätzen zu beschäftigen. „Whole-heartedness“ versteht Dewey als Enthusiasmus, sich intensiv mit Problemen auseinanderzusetzen und eine Sache tiefer zu ergründen und zuletzt „responsibility“ als Verantwortungsübernahme vom Anfang eines Prozesses beginnend bis zu den daraus resultierenden Konsequenzen. Die dritte Komponente ist der Bereich Sprache und Kommunikation. Diese beinhaltet sprachlichen Prozesse, in denen Denkprozesse eingebunden sind, und Gedanken zum Ausdruck kommen. Gemeint sind soziale Prozesse, die eine bestimmte kommunikative Situation oder einen Kontext voraussetzen.

Als Grundlage für eine aufgeschlossene Haltung bei Lehrpersonen erachten Harrington, Quinn-Leering und Hodson (1996) ein Zusammenspiel der Eigenschaften Aufgeschlossenheit, Pflichtbewusstheit sowie Ernsthaftigkeit und orientieren sich hierbei an den Ausführungen von Dewey (1933). Lehrpersonen fokussieren so über das eigene Klassenzimmer hinaus und beziehen soziale sowie politische Kontexte in ihre Reflexionen mit ein (Wyss, 2013) (Tabelle 9).

Tabelle 9: Modell von Harrington et al. (1996)

„open-mindedness“	Andere Perspektiven erkennen und anerkennen können
„responsibility“	Moralische und ethische Konsequenzen von Entscheidungen berücksichtigen
„whole-heartedness“	Eigene Probleme erkennen und klären

Harrington et al. (1996) erkennen im Konzept „open-mindedness“ drei Perspektiven. Die Lehrerperspektive (teacher-focused pattern), wobei das Hauptaugenmerk der Reflexionen auf dem Bedürfnis nach Kontrolle der Schülerinnen und Schüler sowie dem Klassenzimmermanagement liegt. In der Schhülerperspektive (child-focused pattern) fokussieren die Lehrpersonen die Bedürfnisse der Lernenden. Im Vergleich zu der Lehrerperspektive sind hier die Einstellungen wesentlich offener für Kritik. Während die ersten beiden Konzepte sich auf Reflexionen über andere und gekoppelte Verpflichtungen beziehen, beinhaltet das dritte Konzept, die Persönlichkeitsperspektive (inclusive pattern), die Selbstreflexion über persönliche Voraussetzungen (Wyss, 2013).

Nach Henderson (1992) sind Lehrer und Lehrerinnen mit hoher Reflexionskompetenz Experten in dem jeweiligen Unterrichtsfach und im Unterrichten selbst. Dies beinhaltet Fähigkeiten wie gutes Zeitmanagement, Umsetzung von Disziplin, die richtige Auswahl verschiedener Unterrichtsmethoden, gute zwischenmenschliche Beziehungen und das Beherrschen aktueller fachdidaktischer Theorien. Diese Lehrpersonen sind sich der Verantwortung, Entscheidungen treffen zu können, bewusst und nutzen dies als Lerngelegenheit. Ebenso reflektieren sie regelmäßig über die daraus folgenden Konsequenzen. Sie sind für neuen Forschungsergebnisse offen und lernen regelmäßig von ihren reflexiven Erfahrungen. Henderson (1992) erwähnt hier drei wesentliche Merkmale:

1. Hohe Sozialkompetenz
2. Konstruktivistische Einstellung zum Unterricht
3. Hohe Problemlösekompetenz

Zeichner und Liston (1996, S. 6) fanden in ihrem Modell der reflektierenden Lehrperson sechs Dimensionen:

1. „Die Liebe zum Unterrichten
2. Grundlagenwissen
3. Sozialkompetenz
4. Konstruktivistische Einstellung zum Unterricht
5. Hohe Problemlösekompetenz
6. Forschende Einstellung zum Unterrichten“

Dymoke und Harrison (2008) arbeiteten fünf Kernkompetenzen für reflektierendes Denken heraus: Beobachtung, Kommunikation, Urteilsvermögen, Entscheidungsfindung und Teamarbeit. Alle fünf Punkte des reflektierenden Denkens sind wichtige berufliche Fähigkeiten und gelten für alle Phasen des Lehrerlernens, einschließlich der ersten Lehrerausbildungsphase (Mirzaei, Phang & Kashefi, 2014).

Mönig (2013) erarbeitete in ihrer Doktorarbeit drei Dimensionen wissenschaftlicher Reflexionsfähigkeit (Abbildung 15):

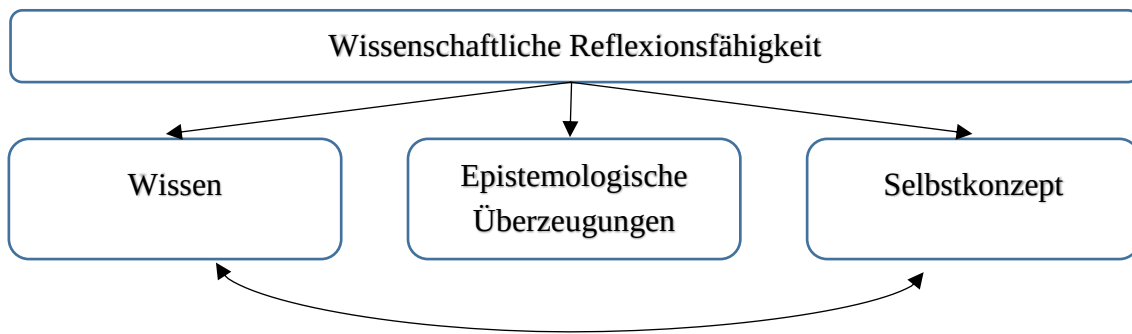


Abbildung 15: Basisdimensionen wissenschaftlicher Reflexionsfähigkeit (Mönig, 2013, Hierbei unterteilt sie die Komponente Wissen in die Teilfähigkeiten Forschungsergebnisse, Studien, Wissenschaftssprache, Forschungslogik, Methoden, Zitation und statistische Grundbegriffe. Epistemologische Überzeugungen beinhalten Sicherheit von Wissen, Umgang mit Autoritäten, reflexive Natur von Wissen, soziale Komponenten, Wertigkeit von Wissen und Lernen lernen. Die letzte Komponente, das Selbstkonzept, steht für sich und wird nicht weiter unterteilt. Hinsichtlich der Professionalität angehender Lehrender wird das Prozesshafte, die Einsicht in die Notwendigkeit zur ständigen Weiterentwicklung und zu lebenslangem Lernen des Öfteren betont (z. B. Nittel, 2011; Helsper & Timpelt, 2011). Es ist die Überzeugung notwendig, dass Wissen stets relativ ist und nie endgültig sein kann und deswegen eine Weiterentwicklung unabdingbar ist. Mönig (2013) ist der Meinung, dass hierbei epistemologische, konstruktivistisch geprägte Überzeugungen förderlich sind. Sie folgert im Weiteren, dass es aufgrund des schnellen Wandels von Wissensständen und der Vielzahl an unterschiedlichen Sichtweisen die Aufgabe der Lehrenden ist, selbstständig und kritisch die verschiedenen Lösungsansätze heranzuziehen und reflektiert in der Praxis anzuwenden. Im Rahmen ihrer Arbeit geht sie davon aus, dass das wissenschaftsbezogene Selbstkonzept und die Fähigkeit zum wissenschaftlichen Reflektieren in Zusammenhang stehen.

Es zeigt sich also insgesamt, dass die Reflexionsprofession im Lehrberuf auf Wissen im Bereich der erziehungswissenschaftlichen Forschung basiert sowie auf der Einsicht in die Notwendigkeit der Anwendung der wissenschaftlichen Erkenntnisse in der pädagogischen Praxis. Auch motivationale Aspekte werden immer wieder angeführt. So kristallisieren sich die drei Elemente des Wissens, der Überzeugungen und der Motivationen als Grundlage des Modells der wissenschaftlichen Reflexionsfähigkeit heraus (Mönig, 2013).

3.3. Förderung der Reflexionskompetenz

In diesem Kapitel sollen verschiedene Ansätze zur Förderung der Reflexionskompetenz dargestellt werden. Um zunächst eine geeignete Umgebung für eine erfolgreiche Einleitung des Reflexionsprozesses zu schaffen, schlägt Schüpbach (2007) vor, ausreichend Zeit einzuplanen, Vertrauen zu schaffen und ritualisierte Abläufe einzuhalten. Korthagen und Kessels (1999) fordern in ihrem Beitrag, die Pädagogik der Lehrerbildung zu überarbeiten und den Studierenden bei der Reflexion mehr Unterstützung zukommen zu lassen. So ergänzten sie das ALACT-Modell (Kapitel 3.1.2) um einen äußeren Kreis der Förderung des Reflexionsprozesses (Abbildung 16):

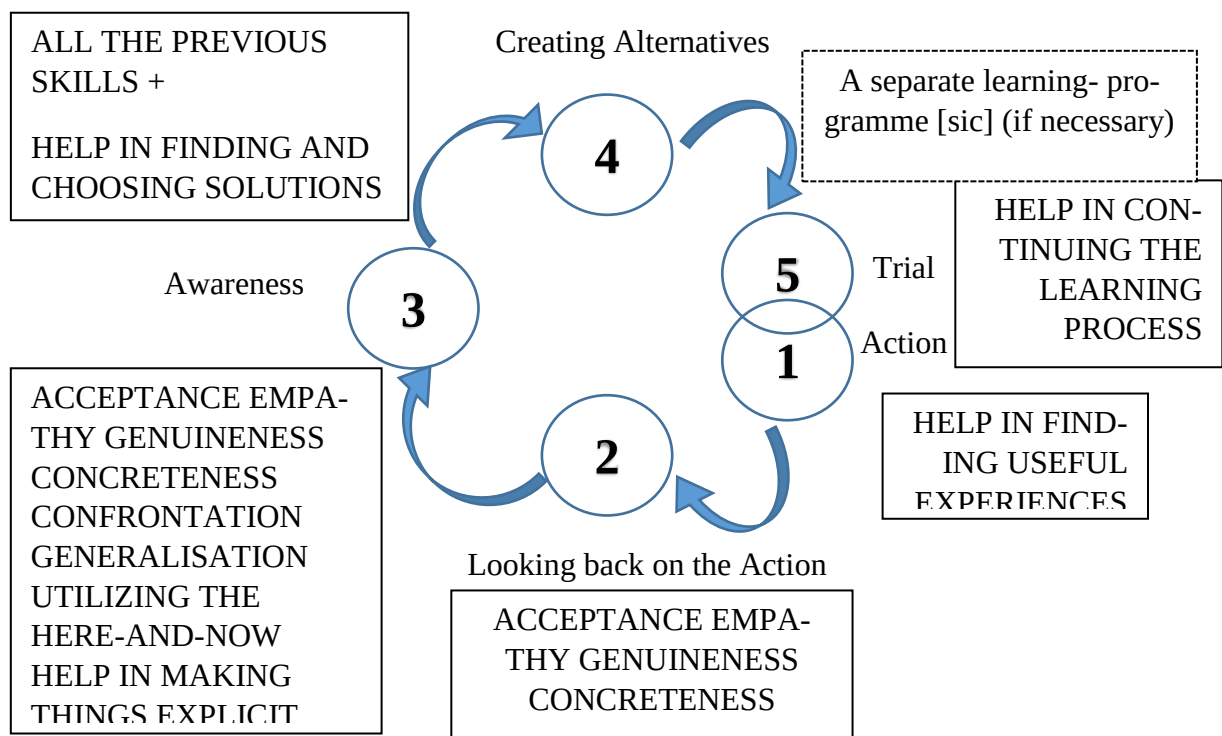


Abbildung 16: Das ALACT-Modell und die nötige Lehrerausbildungskompetenz (Korthagen & Kessels, 1999, S. 13)

Nach Dymoke und Harrison (2008) kann reflektierendes Denken auf verschiedene Arten unterstützt werden. Sie nennen Videoaufzeichnungen, schriftliche Aufsätze, Zeichnungen, Fotografien, Portfolio, Co-Teaching und die Aktionsforschung. Im Folgenden werden die in dieser Studie gewählten Ansätze näher erläutert.

3.3.1. Lehr-Lern-Labor

Seit den 1990ern entwickelte sich das Format der Lehr-Lern-Labore in Deutschland stetig weiter. Diese besitzen zentrale Merkmale, wie das selbstständige Lernen und den Wissenserwerb der Schülerinnen und Schüler und auch der Lehramtsstudierenden. Dies alles geschieht in einer Umgebung, welche als konstruktiv und lernprozessaktivierend wahrgenommen wird. Als weiteren Vorteil sieht Münzinger (2001) den Einsatz experimentell-forschender Methoden zur Auswertung und zur Anregung des Austausches zwischen Lehrerinnen und Lehrern, Studierenden sowie Wissenschaftlerinnen und Wissenschaftlern. Ebenso betont er die Möglichkeit, aus misslungenen Lehrversuchen zu lernen und die daraus resultierenden, gewinnbringenden Bearbeitungsmöglichkeiten und Reflexionen. Auch der Praxisschock kann, durch das Üben des Unterrichtens, welches in seiner Komplexität etwas reduziert ist, abgemildert werden (Fischler, 2008). Einen interessanten Ansatz für zyklisch forschendes Lernen im Lehr-Lern-Labor schlagen Nordmeier et al. (2014) vor (Abbildung 17):

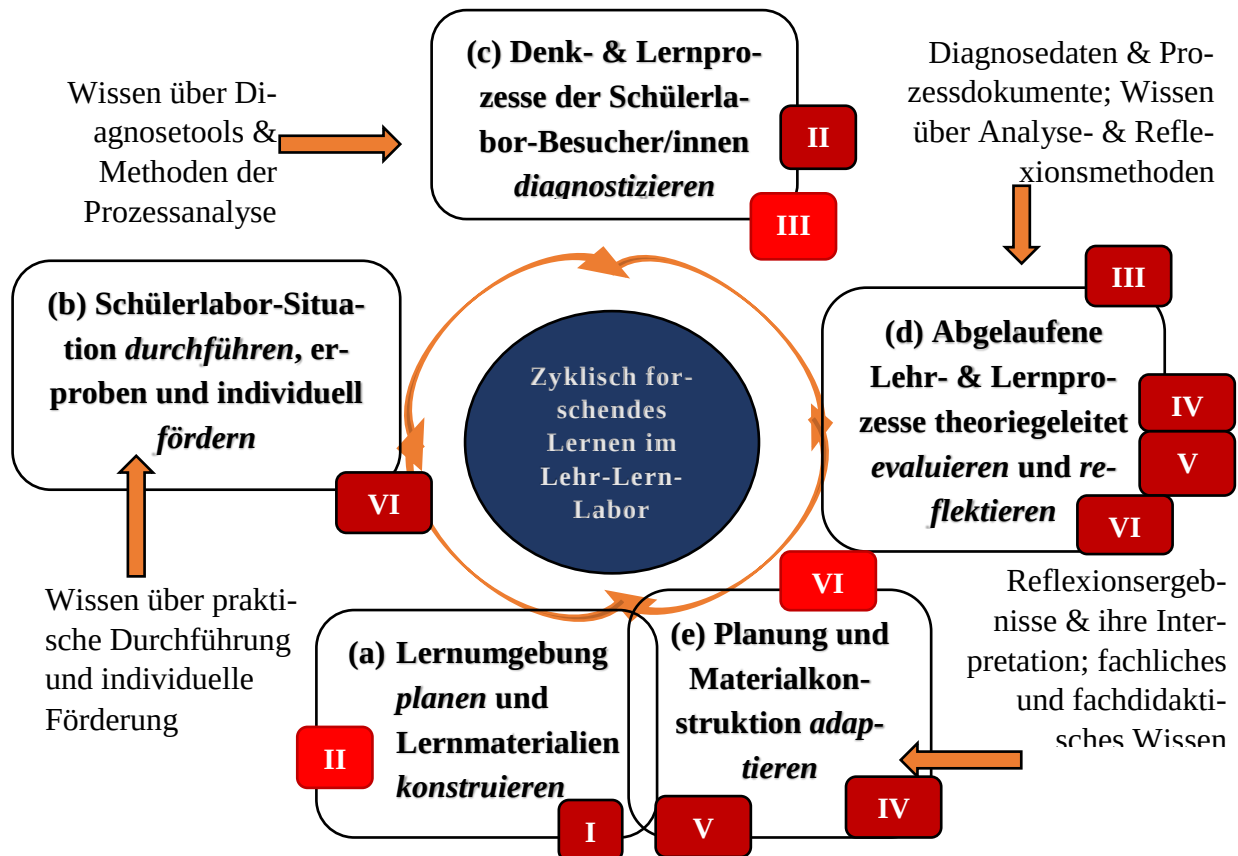


Abbildung 17: Zyklisch forschendes Lernen im Lehr-Lern-Labor (idealisiert) (Nordmeier et al., 2014)

Rehfeldt et al. (2018, S. 104) ergänzen diesen Zyklus um das „Kompetenzmodell zur professionellen Wahrnehmung und Handeln [sic] im Unterricht“ von Barth (2017, S. 40). Die einzelnen Stufen sind hierbei wie folgt benannt:

- I: Wissen
- II: Erkennen
- III: Beurteilen
- IV: Generieren
- V: Entscheiden
- VI: Implementieren

Rehfeldt et al. (2018) ordnen dabei dem Schritt (a) die Facette Wissen zu, da zunächst theoretische Grundlagen des Fachs und der Fachdidaktik gelegt werden. Hier werden erste beispielhafte Unterrichtsvideos eingesetzt, sodass die Facette Erkennen ebenso zugeordnet werden kann. Bei der Durchführung (b) unterscheiden sie in unterrichtende Lehramtsstudierende, Facette Implementation, und aktiv beobachtende Mitstudierende, Facette Erkennen und Beurteilen. Eine besondere Bedeutung geben sie dem Reflexionsschritt (d), da hier vier von sechs Facetten zugeteilt werden können. Situationen werden beurteilt, Handlungsalternativen erarbeitet, Entscheidungen gefällt und die Reflexion der Handlung vorgenommen. Im letzten Adaptionsschritt (e) wird generiert und entschieden sowie implementiert.

Die Ergebnisse der Studie von Rehfeldt, Klempin und Nordmeier (2018) sind für die Entwicklung der Reflexionskompetenz in Lehr-Lern-Laboren umfassend. Um die Reflexionstiefe erfassen zu können, wurden Daten über schriftliche, offene Diskursvignetten

zum fachspezifischen, fachdidaktischen Schwerpunkt im Pre-Postdesign ermittelt. Bei einer Stichprobe von $N = 81$ ($N_{LLS} = 50$, $N_{Control} = 31$) zeigte sich sowohl bei dem Seminar, welches in Lehr-Lern-Laboren unterrichtete ($M_{PRE} = 2.28(0.09)$, $M_{POST} = 3.26(0.10)$), als auch bei der Kontrollgruppe ($M_{PRE} = 1.97(0.09)$, $M_{POST} = 2.32(0.14)$) ein Zuwachs an Reflexionstiefe. Jedoch konnte das Lehr-Lern-Labor-Seminar im Mittel einen höheren Zuwachs mit $d = 1.07$ [CI: 0.65; 1.50] verbuchen als das Theorieseminar der Kontrollgruppe ($d = 0.50$ [CI: 0.02; 1.02]).

3.3.2. Videografie

Potentiale von Unterrichtsvideos

Um Theorie und Praxis zu verbinden und um das anwendbare Wissen bei Lehramtsstudierenden zu steigern, wird von Forschungsseite empfohlen, Lernsettings zu schaffen, welche den Übergang erleichtern (Blomberg, 2011; Brouwer & Korthagen, 2005; Darling-Hammond, 2012; Korthagen & Kessels, 1999; Rovegno, 1993). Einen vielversprechenden Ansatz liefert hier der Einsatz von videografierten Unterrichtsstunden, der es ermöglicht, erste praktische Erfahrungen bereits in der universitären Lehrerbildung zu erlangen und so theoretisches Wissen anzuwenden (Barron, 2007; Blomberg, 2011; Brophy, 2003). Der Praxisschock lässt sich somit abmildern, da die systematische Betrachtung der Unterrichtsvideos erlaubt, theoretisch Gelerntes an Beispielen praktisch anzuwenden sowie zu reflektieren (Mayr, Eder & Fartacek, 1988). Ebenso bietet sich die Möglichkeit, die einzelnen Phasen einer Reflexion anhand einer Grundlage zu bearbeiten, was insbesondere Berufsanfängern, welche die komplexen Anforderungen noch nicht einschätzen können, als Stütze dient. Auf diese Weise können sie ein Bewusstsein für die tatsächlichen Herausforderungen entwickeln (Mietzel, 2017). So wird bei neueren Ansätzen in der Lehrerbildungsforschung auf fallbasiertes, forschendes Reflektieren und Analysieren von Unterricht abgezielt, um ein tieferes Verständnis von Lehr-Lernprozessen zu erreichen (Krammer & Reusser, 2005). Weitere Vorteile nach Krammer und Reusser (2005) sind die Möglichkeiten, subjektive Theorien über Unterrichtsprozesse zu erweitern, die Flexibilität im Nachdenken über alternative didaktische Handlungsmöglichkeiten zu steigern, der Aufbau einer gemeinsamen Berufssprache und die Integration verschiedener Perspektiven im Hinblick auf Unterricht. Ebenso ist ein wiederholtes Abspielen möglich, sodass kein unmittelbarer Handlungsdruck besteht (Krammer et al., 2010). Zusätzlich kann auch nonverbales Verhalten wie Gestik und Mimik analysiert werden (Hosenfeld, 2010).

Vorbehalte gegenüber dem Einsatz von Unterrichtsvideos

Neben einem erhöhten organisatorischen Aufwand, man bedenke technische, zeitliche und finanzielle Aspekte, muss zusätzlich der immer komplizierter werdende Datenschutz beachtet werden (Tuma, Schnettler & Knoblauch, 2013). Ebenso kann, abhängig vom Fokus der Kamera, immer nur ein gewisser Ausschnitt der Realität gezeigt werden. Hierbei wäre es hilfreich, möglichst viele Zusatzinformationen bereitzustellen, was sich wiederum nur schwer umsetzen lässt. Aufgrund der fehlenden Option der Interaktion mit den Akteuren im Video bleibt die Wahrnehmung, ebenso wie die Interpretation der Unterrichtssituationen, subjektiv (Krammer et al., 2010). Seago (2003) gibt zusätzlich zu bedenken, dass Lehrpersonen vorschnell Urteile über die Qualität des Unterrichts abgeben, ohne dabei Bezüge zu Denkweisen und Lernprozessen der Schülerinnen und Schüler herzustellen. In bisherigen Studien (Borko, Jacobs, Eiteljorg & Pittman, 2008; Gärtner, 2007;

Sherin, 2007) führte die Beobachtung allein somit nicht zur Erweiterung des Professionswissens. Hierbei ist eine sorgfältige Gestaltung des Lernsettings unabdingbar. Vorgesprochen werden Möglichkeiten zum gemeinsamen Austausch oder eine strukturierte Anleitung bzw. Begleitung der Arbeit mit den Videos. Ebenso relevant für den erfolgreichen Einsatz haben sich ein gemeinsamer Austausch über den Unterricht und die Etablierung einer konstruktiven Reflexion und Diskussion erwiesen (Krammer & Reusser, 2005).

Formen des Einsatzes

Es werden zwei Grundfunktionen des Einsatzes von Videos unterschieden (Pauli & Reusser, 2006; Seago, 2003): Zum einen können erwünschte Verhaltensweisen mit dem Ziel der Aneignung gezeigt und illustriert werden und zum anderen ist es möglich, über gezeigte Unterrichtseinheiten zu reflektieren, mit dem Ziel einer verbesserten Wahrnehmung der Qualität von Lehr-Lernprozessen (Krammer & Reusser, 2005). Bei Letzterem bietet sich zudem die Möglichkeit, eigenes Videomaterial selbst zu reflektieren (Abbildung 18):

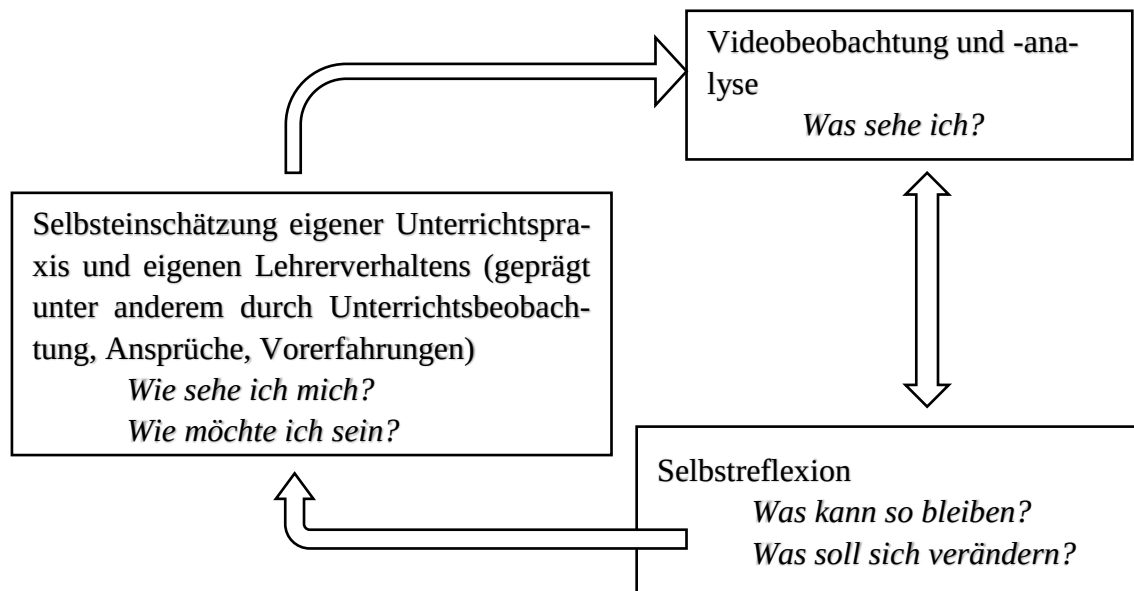


Abbildung 18: Selbstreflexion anhand eigener Videoaufzeichnungen (Nitsche, 2014, S. 87)

Nitsche (2014) nimmt an, dass sich die Prozesse Beobachtung, Analyse und Selbstreflexion nicht immer klar voneinander unterscheiden lassen. So könnten sich bereits während der Videobeobachtung Antworten auf Reflexionsfragen ergeben. Ebenso schließt U. Mühlhausen (2012) nicht aus, dass es bei der Verwendung von Eigenvideos zu einer Fokussierung auf die Person des Unterrichtenden kommt und somit der Blick für bestimmte Aspekte des Unterrichts, wie strukturelle oder konzeptionelle Punkte, erschwert wird. Um dieses Problem zu umgehen, plädiert er für die Verwendung von sowohl Eigen- als auch Fremdvideos. Bei der reinen Reflexion von Aufzeichnungen fremden Unterrichts entsteht der Nachteil, dass die Voraussetzungen, unter denen der geplante Unterricht entstanden ist, nicht rekonstruierbar sind. Gemeint sind zum Beispiel das didaktische Konzept, die Unterrichtsplanung, die Qualität der verwendeten Materialien und die Resultate, welche sich nur unzureichend mittels der Videos erkennen lassen (Mühlhausen, 2005). Um dem entgegenzuwirken, könnte man die Videoaufzeichnungen durch Begleitmaterialien ergänzen. Die Ergebnisse des Projekts „Lernen aus Unterrichtsvideos“ (LUV) bleiben nach

eigenen Angaben der Autorin hinter den Erwartungen zurück und verorten die Wahrnehmung vornehmlich auf beschreibendem Niveau (Schwindt, 2008). Als Verbesserungsvorschlag wird die Beobachtung von Unterrichtsvideos empfohlen, welche den persönlichen Erfahrungen und dem Ausbildungsstand entsprechen. So sind nach Brophy (2003) diejenigen Videos die besten, welche Lehrpersonen zeigen, mit denen die Zuschauer sich identifizieren können, zum Beispiel durch Umsetzung eines ähnlichen Lehrplans oder durch Unterrichten in einem Klassenzimmer, das dem, in dem sie selbst unterrichten, ähnlich ist oder sein wird. So müssen Situationen geschaffen werden, in denen die persönliche Übertragung funktioniert (Nitsche, 2014). Reusser (2005) sieht neben dem videobasierten Lernen am Modell, nach dem Lernen mithilfe der Beobachtung vorbildlichen Lehrerinnen- und Lehrerverhaltens angestrebt wird, und der problemorientierten und fallbasierten Analyse von Unterrichtsvideos, bei der unterschiedliche Perspektiven betrachtet werden, die videogestützte Reflexion und Feedback als weitere Einsatzmöglichkeiten. Er schreibt: „Reflexives, feedbackorientiertes Lernen [...] anhand von – vorzugsweise eigenem – videografiertem Unterricht zielt auf die bewusste Wahrnehmung (Selbstkonfrontation) und Vergegenwärtigung des eigenen Unterrichtshandelns aus einer Außenperspektive“ (Reusser, 2005, S. 12). Ausbildungsziele sieht er im Bewusstwerden von schlechten Angewohnheiten und im Erkennen von Handlungsstrategien und Theorien über Lehrpersonen und Lernen zur Konstruktion von Gestaltungsalternativen.

Lernen durch Reflexion und Analyse von Unterrichtsvideos

Die auf den ersten Blick ähnlichen Begriffe Analyse und Reflexion lassen sich wie folgt unterscheiden. So ist die Analyse nach Mühlhausen (2012) eine kriteriengeleitete Untersuchung mit Distanzierung vom Analysegegenstand. Die Reflexion hingegen ist inhaltlich frei, mit dem Ziel, Handlungsalternativen zu entwickeln. Bei der Beschäftigung mit dem eigenen Verhalten im Unterricht wird ein Kompetenzausbau ermöglicht. Nach Mayer-Frühwirth (2017) gelang es den Lehrerinnen und Lehrern nicht, sich während der Unterrichtssituation auf bestimmte Aspekte zu fokussieren; hierbei nehmen Mühlhausen (2012) Bezug auf Schöns Theorie zur Reflexion in Handlungssituationen (reflection in action) (Kapitel 3.1.2). Sie fanden vorwiegend Beobachtungen auf bereits existierenden Grundannahmen und impliziten Theorien. Nur durch das wiederholte Betrachten des eigenen Videomaterials wurde es den Lehrerinnen und Lehrern erleichtert, unbemerkte Situationen zu analysieren und zu reflektieren (Krammer & Reusser, 2005). Dabei bietet die videografierte Unterrichtsstunde die Möglichkeit der Reflexion einer vorausgegangenen Handlung im Nachhinein (reflection on action) nach Schön (Mayer-Frühwirth, 2017). Krammer und Reusser (2005) unterscheiden zwei Lernarten bei der Reflexion von Unterrichtsvideos, das fallbasierte, problemorientierte Lernen und das forschende Lernen. Ersteres bietet die Möglichkeit von ko-konstruktivem Lernen. Das heißt, es findet ein gemeinsamer Austausch statt und eine bewusste Einnahme verschiedener Perspektiven. Um dies besser nachvollziehen zu können, werden zusätzlich Begleitmaterialien wie Arbeitsblätter oder Transkripte zu Verfügung gestellt und die Unterrichtsreflexion angeleitet. So sollen u. a. längerfristige professionelle Entwicklungsprozesse angeleitet werden (Le Fevre, 2003). Das forschende Lernen verfolgt das Ziel der professionellen Kommunikation über Unterrichtsprozesse, die Fokussierung auf Qualitätsmerkmale guten Unterrichts und eine strukturierte Analyse von Lehr-Lernprozessen. Hier sei auf die Konzepte der Aktionsforschung (Kapitel 3.1.2) verwiesen. Wesentliche Merkmale sind der enge Bezug von Aktion und Reflexion um neue Möglichkeiten zu schaffen, der Praxisbezug

sowie längerfristige Forschungs- und Entwicklungszyklen mit der Einnahme unterschiedlicher Perspektiven. Altrichter et al. (2018) inkludieren zusätzlich den Austausch mit Kolleginnen und Kollegen, Wissenschaftlerinnen und Wissenschaftlern oder Lehrerfortbildnern.

Herausforderungen bei der Reflexion und Analyse von Unterrichtsvideos

Bei der Reflexion und Analyse von Unterrichtsvideos ist es nötig, den jeweiligen Beobachtern die Subjektivität der eigenen Wahrnehmung bewusst zu machen (Krammer & Reusser, 2005). So können Effekte wie der Primacy-Effekt (erster Eindruck prägt weitere Beobachtungen), der Halo-Effekt (eine einzelne Beobachtung überdeckt andere Situationen), die Stereotypisierung (Vorurteile je nach Schicht, Geschlecht, Ethnie, ...) und die Projektion (eigene Mängel oder Bedürfnisse werden auf die betrachtete Person übertragen) gemildert werden (Stroebe, Hewstone & Stephenson, 1996). Ebenso gilt es, die Form des Diskurses zu beachten (Krammer & Reusser, 2005). Seago (2003) konnte zwei vermehrt auftretende Aussagearten von Lehrpersonen bei der Reflexion der Videos nachweisen: Zum einen wurde sehr schnell in gut und schlecht eingeteilt bzw. als richtig und falsch bewertet. Und zum anderen konnte eine Tendenz nachgewiesen werden, nur Positives zu sagen. Auch Krammer und Reusser (2005) haben dies bestätigt. So entstand der Vorschlag, die Reflexion in einem gewissen Maß anzuleiten. Den reflektierenden Lehrpersonen muss bewusst werden, dass nicht allein die Bewertung maßgeblich ist, sondern es wichtig ist, „die Objektivierung und die analytische Durchdringung der beobachteten Lehr-Lernprozesse mit Blick auf deren psychologisch-didaktische Tiefenstruktur“ (Krammer & Reusser, 2005, S. 44) zu erreichen.

3.3.3. Learning Study

Unter dem Begriff Learning Study lässt sich ein Konzept zur Förderung bzw. Unterstützung der Entwicklung von Lehrerinnen und Lehrern verstehen. Es basiert auf dem Ansatz der sogenannten Lesson Study (Cheng & Lo, 2013; Tan & Nashon, 2013). Letztere besteht aus folgenden Prozessen: Zunächst wird eine Unterrichtsstunde entwickelt, welche im Anschluss in einer Übungsstunde erprobt wird. Es folgt die Reflexion der gehaltenen Stunde. Die gewonnenen Erkenntnisse werden in den täglichen Unterricht eingebracht (Kuno, 2011). Bei der Reflexion wird unter anderem die Videoreflexion als eines der effektivsten Hilfsmittel genannt. Nach Lewis, Perry und Hurd (2004) konnten sich japanische Lehrerinnen und Lehrer durch die Lesson Study im Fachwissen verbessern, mehr Wissen über den Unterricht ansammeln, ihre Beobachtungsgabe steigern, stärkere kollegiale Netzwerke aufbauen, sich besser mit langfristigen Zielen identifizieren, besser motivieren und eine höhere Qualität der verfügbaren Unterrichtspläne gewährleisten. Es zeigten sich vier wesentliche Vorteile (Cajkler, Wood, Norton, Pedder & Xu, 2015, S. 194):

- Bessere Zusammenarbeit der Lehrerinnen und Lehrer
- Schärferer Fokus auf den Lernfortschritt der Schülerinnen und Schüler
- Entwicklung des Fachwissens, der Lehrpraxis und der Professionalität
- Verbesserte Qualität des Unterrichts und der Lernergebnisse der Schülerinnen und Schüler

Im Gegensatz dazu beziehen sich Learning Studies nicht ausschließlich auf die Analyse und Weiterentwicklung von Unterrichtsstunden, sondern auch auf Schüleraktivierung,

mit dem Ziel, die Art und Weise des Lernens der Schülerinnen und Schüler zu erlernen (Wood & Sithamparam, 2014). Eine weitere Unterscheidung treffen Tan und Nashon (2013), indem sie dem Prinzip der Learning Study die Verwendung eines theoretischen Rahmens zuordnen (Holmqvist, 2011; Pang & Lo Ling, 2012). Laut Pang und Marton (2003) werden instrumentelle und theoretische Ansätze der Lehrerbildungsforschung kombiniert, um den Mangel des theoretischen Rahmens der Lesson Studies auszugleichen. Darüber hinaus kann der Bezug zur Theorie den Diskurs vervielfältigen. So können sich die Lehrpersonen beispielsweise in einer Unterrichtsstunde auf Klassenaktivitäten und Unterrichtsmaterialien konzentrieren (Fernandez, Cannon & Chokshi, 2003). Mit Hilfe einer Learning Study lassen sich Schülerlernen und Lehrerhandlung stattdessen aus einer theoretischen Perspektive interpretieren und diskutieren (Pang & Lo Ling, 2012). Es bleibt zu betonen, dass die Planung einer Unterrichtsstunde eine komplexe Aufgabe ist, welche unterschiedlichste Kapazitäten über mehrere Zeiträume hinweg erfordert. Dies kann sich auf Zeiträume innerhalb einer Unterrichtsstunde oder ganze Wochen beziehen (Calderhead, 1996). Folglich gibt es unterschiedliche Ansätze. So definieren Pang und Marton (2005) das Prinzip der Learning Study als einen professionellen Entwicklungsansatz, der Lehrpersonen ermutigt, Lernaktivitäten von Schülerinnen und Schülern gemeinsam zu planen, durchzuführen und zu bewerten. Nilsson (2014) versteht darunter einen kollegialen Prozess, in welchem Lehrerinnen und Lehrer mit einem Forscher zusammenarbeiten, um ihre eigenen Lehreraktivitäten zu erforschen mit dem Ziel, zu ermitteln, was für das Lernen ihrer Schülerinnen und Schüler entscheidend ist. Hierbei wird das eigene Verhalten im Unterricht erkundet, reflektiert und in Beziehung zum Lernzuwachs der jeweiligen Schülerinnen und Schüler gesetzt. Ähnlich zur Aktionsforschung agiert die Forscherin bzw. der Forscher als eine Art Tutor, welcher die Teilnehmerinnen und Teilnehmer in ihren Reflexionsphasen unterstützt (Runesson, 2008). Runesson (2008) betont, dass das Prinzip mehr ist als nur reflexive Praxis. Sie beschreibt es als einen systematischen Prozess, der auf einem theoretischen Rahmen basiert. Im Fokus steht nicht die Stunde als solche, sondern die zahlreichen Lernmöglichkeiten der Studierenden. Der Ablauf unterteilt sich in sechs Hauptphasen (Abbildung 19):

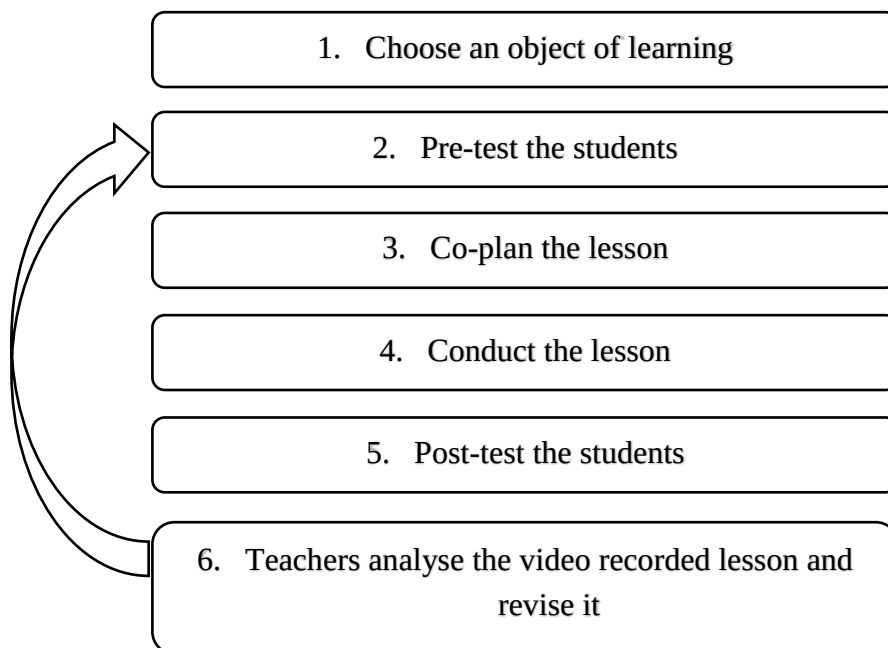


Abbildung 19: Ablauf einer Learning Study (Nilsson, 2014, S. 1799)

Die Lehrpersonen wählen ein Lernobjekt aus, welches die Schülerinnen und Schüler untersuchen sollen. Ein gut gewähltes Objekt wird demnach zu Erklärungen führen, die inhaltliches Denken erfordern und entwickeln. Dieses Prinzip lässt also Lehrerinnen und Lehrer komplexe Aufgaben konstruieren, welche wiederum Schülerinnen und Schülern zu problemlösendem Denken anregen. Die Erhebung des Vorwissens der Schülerinnen und Schüler wird ausgewertet und etwaige Wissenslücken oder Fehlvorstellungen an die Lehrerinnen und Lehrer weitergegeben. Aufgabe der Lehrpersonen ist es, das Denken der Schülerinnen und Schüler zu antizipieren und zu analysieren. Auf diese Weise veranschaulicht das Prinzip, was Lehrerinnen und Lehrer wissen und was sie über ihre Schülerinnen und Schüler lernen. Es folgt in Zusammenarbeit mit einem Tutor die Entwicklung eines Unterrichtskonzepts. Die gehaltene Unterrichtsstunde wird schließlich videografiert und umfassend analysiert. Direkt im Anschluss der Stunde werden die Schülerinnen und Schüler erneut getestet. Mithilfe dieser und der restlichen ausgewerteten Ergebnisse wird die Stunde überarbeitet und weiterentwickelt (Nilsson, 2014). Hierbei kann durch geschickte Einbettung des fachdidaktischen Wissens (PCK) in die Aufgabenstellung für die Lehrerinnen und Lehrer die Entwicklung des PCKs veranschaulicht werden (Schneider, 2019). So schreibt Schneider (2019) weiter, dass Learning Study und PCK direkt mit der Praxis des Lehrens verbunden sind und folgern, dass Lehrerinnen und Lehrer mittels der Anwendung des Prinzips ihr PCK nutzen und weiterentwickeln werden. Die sich beliebig oft wiederholende Reflexionsphase und der Vergleich der Unterrichtskonzepte können variieren und hängen von der Anzahl der Teilnehmerinnen und Teilnehmer ab.

Studien zum Prinzip der Learning Study betonen, besonders auf die exponierte Stellung des Forschers bzw. Tutors zu achten. Adamson und Walker (2011) fanden in ihrer Studie heraus, dass die Zusammenarbeit von Forschern und Lehrpersonen nicht immer einfach ist. Spannung fanden sie begründet in den unterschiedlich eingenommenen Perspektiven und in unklaren hierarchischen Strukturen. Die Expertise des Forschers unterschied sich immens von der Einschätzung der Lehrerinnen und Lehrer. So waren die Forscher beispielsweise mit dem theoretischen Rahmen besser vertraut, während die Lehrpersonen ein tiefes Verständnis für den Kontext im Klassenzimmer und die Lernenden hatten. Um das Dilemma unterschiedlicher Expertise zu lösen, informierten Runesson und Gustafsson (2012) die Lehrerinnen und Lehrer explizit, dass die Studie in den Händen der Lehrpersonen liegt, wenn es darum geht, die Entscheidungen zu treffen. So wurden die Forscher angehalten, so wenig wie möglich in den von den Lehrerinnen und Lehrern gestalteten und geführten Prozess einzugreifen. Um den Einfluss der Forscher zu minimieren, hat man die Lehrpersonen ermutigt, einen Teil der Planungssitzungen selbstständig durchzuführen.

Olander und Olander (2013) untersuchten in ihrer Studie die Auswirkung der Learning Study auf das fachdidaktische Wissen (PCK) der Lehrerinnen und Lehrer, inwiefern die Entwicklung der Lehrperson auf das Schülerlernen Einfluss hat und in welcher Art und Weise die Entwicklung der Lehrperson in der Unterrichtsgestaltung wiederzufinden ist. Hierbei erarbeiteten Lehrerinnen und Lehrer eine Biologiestunde und durchliefen obigen Kreislauf dreimal (Stunde A, Stunde B, Stunde C). Die jeweils drei Klassen füllten einen Pre- und einen Posttest zum Thema der drei Stunden aus. Die Autoren verzeichneten in allen drei Stunden einen Anstieg der durchschnittlich erreichten Punkte vom Pre- zum Posttest. Der kleinste Lernzuwachs auf Schülerseite war dabei in Stunde A zu erkennen mit einem Durchschnitt von 3.11 Punkten hin zu 3.22 Punkten, sukzessive ansteigend über Stunde B bis hin zur dritten Stunde, mit dem größten Anstieg von anfangs 2.75 Punkten

im Pre-Test und schließlich 5.00 Punkten im Posttest. Hier waren jedoch nur die Ergebnisse in Stunde C signifikant ($t = 2.39$; $p = .048$). Weiter konnten Olander und Olander einen Fachwissenszuwachs ebenso wie ein erweitertes Methodenspektrum von Stunde zu Stunde feststellen. Die Autoren schreiben weiter: „On a general level, we argue that the collaborative and iterative working model (learning study) used in this study is an important tool and has potential to scaffold teachers’ professional development“ (Olander & Olander, 2013, S. 210).

3.3.4. Feedback

Feedback ist eine „auf Daten basierende Rückmeldung an eine Person bzw. Personen-Gruppe zu ihrem vorherigen Verhalten“ (Haag & Götz, 2019, S. 14). Die Autoren sehen in dieser Definition mehrere Vorteile. Zunächst wird der Begriff Beurteilung vermieden, welcher Funktionen wie Selektion bzw. Kontrolle beinhaltet. Zudem lassen sich einfache Formen der Zustimmung oder Ablehnung darin verorten, ebenso elaboriertere Ausführungen wie Ursachen von Fehlern oder Verbesserungsvorschläge. Zudem ist es möglich, Feedback nicht nur von einer, sondern auch von mehreren Personen zu erhalten bzw. zu empfangen. Die Funktionen von Feedback lassen sich in vier Hauptkategorien unterteilen (Fengler, 2009; Narciss, 2006):

- Korrekte Antworten bestätigen und verstärken
- Feedback als Informationsquelle zur Korrektur von Fehlern (auf Basis einer positiven Fehlerkultur)
- Feedback als Motivationsquelle
- Feedback als Tutor für die Bearbeitung neuer Aufgaben, zur Erreichung eines vorgegebenen Ziels

Die Wirkung von Feedback lässt sich mithilfe einiger ausgewählter Modelle erklären. Dabei gibt es drei Grundvoraussetzungen für effektives Feedback:

- es ist nötig und gewollt
- es wird zur passenden Zeit gegeben
- man ist in der Lage und bereit, es zu nutzen (Haag & Götz, 2019)

Um das Zusammenspiel von Reflexion und Feedback zu veranschaulichen, entwickelten Luft und Ingham (1955) das sogenannte Johari-Fenster (Tabelle 10):

Tabelle 10: Das Modell des Johari-Fensters

		dem Selbst	
		bekannt	unbekannt
anderen	bekannt	I Öffentliche Person	II Blinder Fleck
	unbekannt	III Privatperson	IV Unbekannte Person

Hierbei finden sich im öffentlichen Bereich (I) die Informationen, welche der Person über sich selbst bewusst sind und ebenso auch anderen Personen bekannt sind. Im geheimen Bereich (III) sind die Informationen lediglich der Person über sich selbst bewusst, anderen Personen sind diese Informationen nicht zugänglich. Im unbekannten Bereich (IV)

sind weder der eigentlichen Person selbst die Informationen bewusst noch anderen außenstehenden Personen. Zum Thema Feedback führt der Bereich des Blinden Flecks (II). Hier sind nun außenstehende Personen im Besitz der Informationen, die Person selbst ist sich jedoch dieser Eigenschaften nicht bewusst. Mithilfe von Feedback wäre es nun möglich, ein Bewusstsein für negative Aspekte zu erlangen und an diesen zu arbeiten. Dieses Modell ist jedoch kein starres Konstrukt, so können Selbstmitteilung und Feedback die Veränderung der Quadranten beeinflussen und dadurch die Größen variieren (Buhren, 2015; Fengler, 2009). Um die Wirkungsweisen von Feedback zu erklären, gibt es verschiedene Ansätze. Als erste lerntheoretische Grundlage findet man den behavioristischen Ansatz, in welchem Feedback als Verstärker betrachtet wird (Haag & Götz, 2019). Positives Feedback soll hierbei richtiges Verhalten festigen, negatives die Wahrscheinlichkeit erhöhen, falsches Verhalten zu mindern. Letzteres führt jedoch nicht automatisch zu richtigem Verhalten. Der kognitive Ansatz löste in den 1970er/80er Jahren diese Denkweise ab. Hier liegt die Betonung auf einer rein informativen Wirkung bzw. Funktion, um sich über bestehende Misskonzepte bewusst zu werden. Je nach Bewertung des Feedbacks durch den Empfänger werden Konsequenzen ergriffen oder das bisherige Verhalten beibehalten. Ziel ist es, ein Gefühl der Kompetenz bzw. der Selbstbestimmung aufzubauen. Feedback sollte somit sachlich auf gezeigtes Verhalten hinweisen und nicht die Person selbst einbeziehen. Im Folgenden werden exemplarisch zwei Rahmenmodelle hierzu vorgestellt:

Narciss (2006) unterteilt Feedback in drei Facetten. So lassen sich Feedbackwirkungen in der Feedback-Funktionsfacette nach den intendierten kognitiven, meta-kognitiven und motivationalen Gesichtspunkten kategorisieren. Es kann also verstärken und bestätigen, motivieren, instruieren, regulieren, etc. Die Feedback-Form ist unterteilt nach Timing, in den Modus und die Adaptivität. Die erste Facette unterscheidet zeitliche Faktoren wie unmittelbares oder verzögertes Feedback (Kulik & Kulik, 1988). Der Modus unterscheidet Formen wie z. B. Peer-Feedback. Die letzte Facette bezieht sich auf den Inhalt. Hierbei unterscheidet man die Komponenten informativ, also zusätzliche Informationen entsprechend der Intention, und evaluativ, sprich Informationen über den Leistungsstand oder die Korrektheit der Lösung. Welche Facette Anwendung findet, ist abhängig von situativen Faktoren wie zum Beispiel Lehr-Lernzielen oder Lernaufgaben. Die Art und Weise, wie die Informationen verarbeitet werden ist, wiederum anhängig von den individuellen Faktoren wie Vorwissen und Lernmotivation (Abbildung 20).

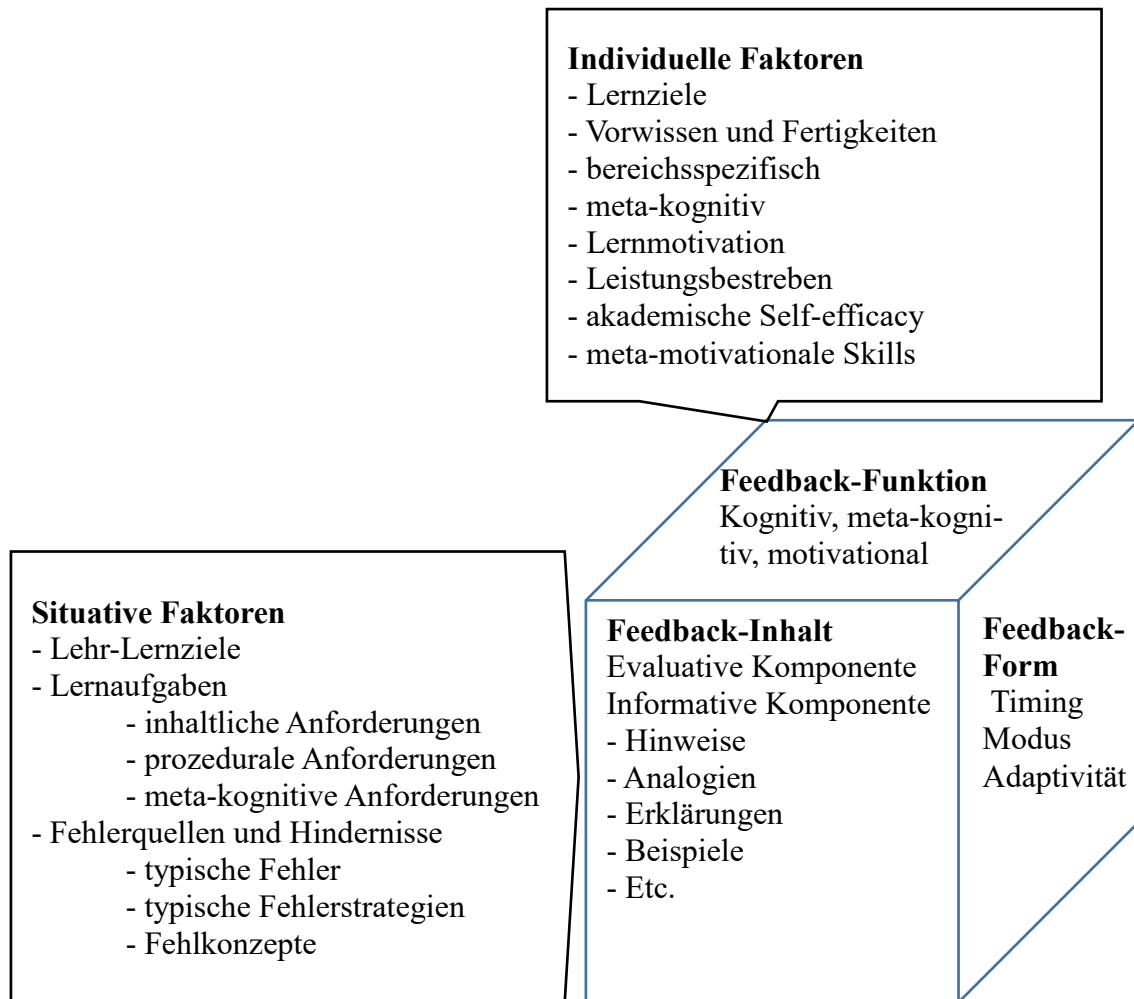


Abbildung 20: Übersicht über die Determinanten des Informationswerts von Feedback (Narciss, 2006, S.81)

Hattie und Timperley (2007) unterscheiden drei zentrale Fragen auf vier unterschiedlichen Ebenen (Abbildung 21).

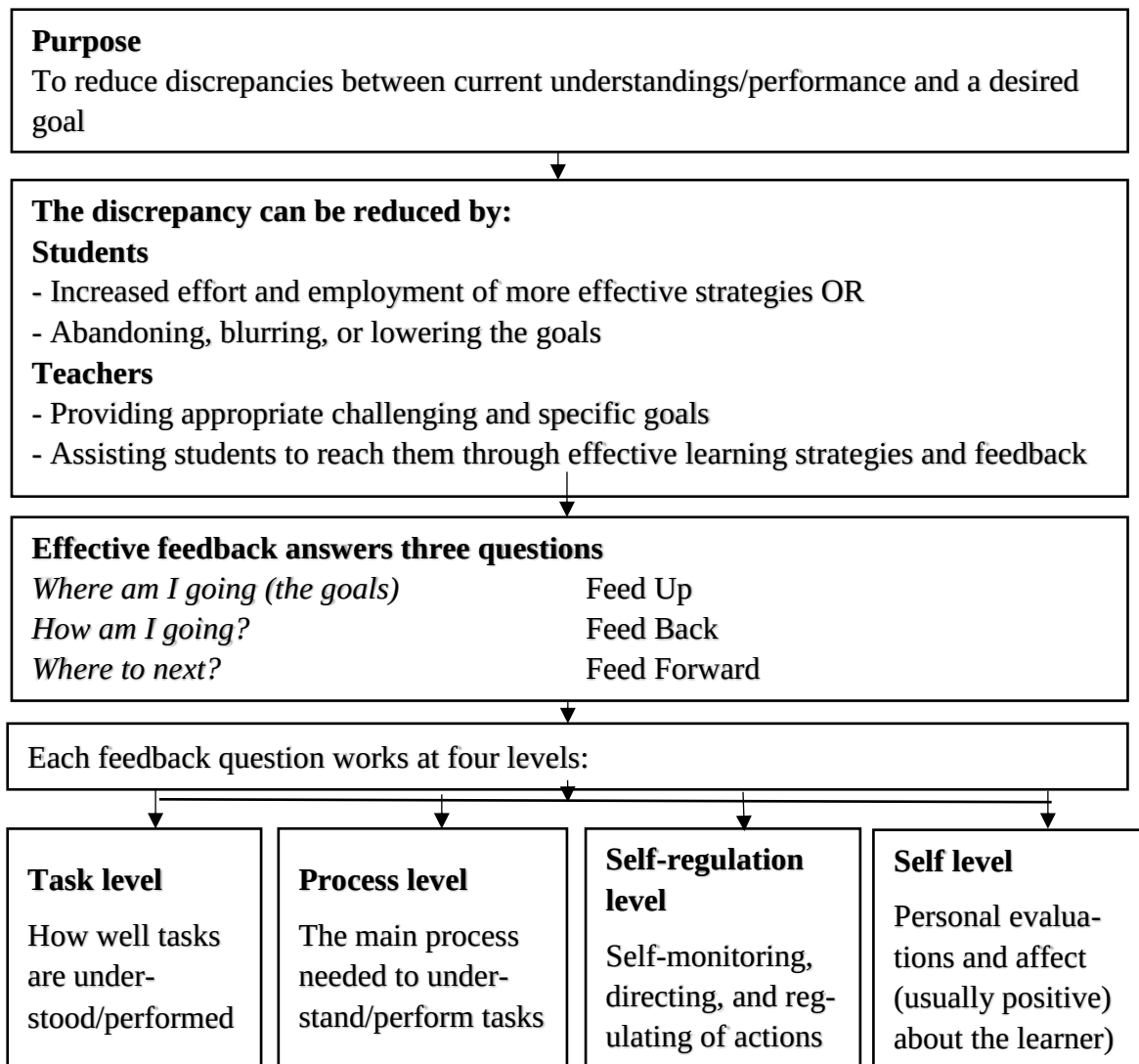


Abbildung 21: A model of feedback to enhance learning (Hattie & Timperley, 2007, S. 87)

Um effektives Feedback zu geben, ist es wichtig, die Umstände zu verstehen, welche zu abweichenden Ergebnissen führen. So müssen drei Hauptfragen beantwortet werden: Was sind die Ziele (Feed Up)? Welche Fortschritte werden auf dem Weg zum Ziel gemacht (Feed Back)? Welche Maßnahmen müssen ergriffen werden, um bessere Fortschritte zu erzielen (Feed Forward)? Wie effektiv Antworten auf diese Fragen sind, ist abhängig von dem Level, auf dem sich der Empfänger befindet. Man unterscheidet die Ebene der Aufgabe, des Prozesses, der Selbstregulation und des Selbst. Ziel der ersten Ebene ist es herauszufinden, wie gut die Aufgaben erledigt wurden. Es entspricht einem Abgleich von Resultat und angestrebtem Ziel. Im Prozess wird eruiert, welche Maßnahmen ergriffen werden müssen, um ein Ziel zu erreichen. Das Selbstregulationslevel analysiert, was der Lernende selbst tun kann, um seine Aufgaben zu verwirklichen. Gemeint sind Planungen, Selbstbeobachtungen, realistische Selbsteinschätzungen und Modifizierung der Strategien (Haase, 2018). Bei der letzten Ebene des Selbst ist Feedback auf die Person selbst bezogen, auf deren Eigenschaften und Merkmale. Hattie (2008) warnt je-

doch vor Feedback auf dieser Ebene, da Persönlichkeitsmerkmale zu stabil sind und negatives Feedback schnell zu einer Abnahme der Motivation bis hin zur Frustration führen kann. Insgesamt wird deutlich, dass sich Feedback nicht nur auf die Korrektheit der Aufgabe bezieht. Effektives Feedback ist kontinuierlich und begleitend zum Lernprozess. Dies wird in der Literatur auch als „formatives Feedback“ bzw. „formative assessment“ bezeichnet (Steffens & Höfer, 2015; Zierer, Busse, Wernke & Otterspeer, 2015). Zudem schreibt Hattie (2012), dass tieferes Lernen Feedback auf der Ebene des Prozesses benötigt. Trotzdem ist ein vollständiges Feedback auf den drei Ebenen die effektivste Art, den größten Lernzuwachs zu erreichen (Hattie & Zierer, 2016).

Die Effektivität von Feedback ist aber noch von weiteren Faktoren abhängig. Mit dem Begriff der Aufgabenart meinen Haag und Götz (2019) vornehmlich den Zeitaspekt. Bei Aufgaben mit deklarativem bzw. prozeduralem Wissen ist unmittelbares Feedback sinnvoll. Bei Transfer- und Konzeptwissen ist hingegen eine aufgeschobene Rückmeldung sinnvoller (Kulik & Kulik, 1988; Shute, 2008). Einen Unterschied macht zudem der Ausbildungsstand der Lernenden. Fortgeschrittene profitieren auch bei leichten Aufgaben von aufgeschobenem Feedback. Im Gegensatz hierzu benötigen gerade Anfänger bei schwierigen Aufgaben unmittelbare Rückmeldung. Über die Häufigkeit sind sich Maier (2010) und Hattie (2008) einig und raten zu einer regelmäßigen Rückmeldung im Gegensatz zu einer summativen am Ende des Zyklus.

Auch für die Feedbackart gibt es einige Befunde. Allgemein gilt, dass Feedback nicht zu lange und zu komplex sein darf, da dies negative Auswirkungen auf die Aufmerksamkeit haben kann (Shute, 2008). Somit gilt als Richtlinie spezifisches und klares Feedback (Haag & Götz, 2019; London, 2003). Weiter ist die Passung zwischen den Zielen und den Erwartungen so zu gestalten, dass sie auch erreicht werden können. Buhren (2015) appelliert, konstruktiv zu bleiben und empfiehlt, konkrete Informationen klar und verständlich zu formulieren. Eine höhere Wirksamkeit als einfache Rückmeldungen weist in der Regel das elaborierte Feedback auf (Bangert-Drowns, Kulik, Kulik & Morgan, 1991; Jaehnig & Miller, 2007). Dies ist jedoch abhängig von der Aufgabenkomplexität und dem Zeitpunkt des Lernens. So ist bei komplexen Aufgaben gepaart mit wenig Vorwissen ein elaboriertes Feedback sinnvoll. Bei Aufgaben mit reinem Faktenwissen genügt oftmals eine Richtig-Falsch-Antwort.

Feedback wirkt also nicht per se. Erklärungen finden Haag und Götz (2019) begründet in mangelndem Vorwissen oder einer nicht ausreichenden Verarbeitungstiefe, was zu einer Zurückweisung bzw. Ignoranz des Feedbacks führen kann. Wenn der Feedback-Empfänger fehlendes Interesse und fehlende metakognitive Voraussetzungen aufweist, werden Informationen nur rezeptiv aufgenommen. Auch wenn Antworten durch Feedback vorweggenommen werden, bevor der Lernende selbst kognitiv aktiv wird, oder wenn die Rückmeldungen nicht zu den kognitiven Bedürfnissen passen, kann dies hinderlich sein.

Feedback gilt als wesentliches Kriterium für gelingende Selbstreflexionen (Jahncke, Berding, Porath & Magh, 2018). Grundsätzlich ist möglich, mittels Beratung und dem Schaffen geeigneter Angebote die Reflexionsfähigkeit zu erweitern (Herzog, 2003; Hilzensauer, 2008). Somit plädieren die in diesem Absatz genannten Autoren für Anleitungen sowie Betreuungs- und Unterstützungsangebote durch Lehrende. Feedback ermöglicht hierbei den Perspektivwechsel und die Aufnahme weiterer unbewusster Aspekte (Bain et al., 1999). Durch regelmäßig wiederholtes Feedback lassen sich die eigenen Erfahrungen

so besser strukturieren und unter Verwendung von Fachbegriffen reflektieren (Hänssig, 2010).

Aktuelle empirische Studien zum Thema Feedback gekoppelt mit Reflexionskompetenz untersuchen die Effekte von Feedbackgesprächen auf die (Selbst-)Reflexionskompetenz. So entwickelten Jahncke et al. (2018) in ihrer Studie ein Seminar, in welchem Studierende Seminarsitzungen in berufs- und wirtschaftspädagogischen Modulen planen, durchführen und reflektieren sollen. Dabei wird die Selbstreflexion von Feedbackgesprächen direkt im Anschluss der Stunde unterstützt. Grundlage der Diskussion ist einerseits ein kriteriengeleiteter Beobachtungsbogen und andererseits eine erste Verschriftlichung zu Leitfragen, welche etwaige Abweichungen von den Planungen oder Gründe und Handlungsalternativen thematisiert. Nach dem Feedback folgt eine zweite Reflexion der Studierenden, um überprüfen zu können, inwiefern die Rückmeldung Einfluss hatte. So sollen Impulse für eine intensivere Auseinandersetzung mit dem Thema gegeben werden. Bei einer Stichprobe von 57 Studierenden wurden die schriftlichen Reflexionen mittels Kodierleitfaden ausgewertet. Dieser unterscheidet die folgenden vier Stufen: deskriptive Darstellung, deskriptive Selbstreflexion, begründete Selbstreflexion und verknüpfende Selbstreflexion. Nach einer reinen Beschreibung der Handlungen werden bei der deskriptiven Selbstreflexion erste Bewertungen vorgenommen. Bei der nächsten Stufe werden zusätzlich Ursachen genannt. Die letzte Stufe stellt darüber hinaus Verbindungen zu Theorien oder vergangenen Erfahrungen her. In der Entwicklungsstudie konnten die Autoren eine Interoderreliabilität gemessen an Krippendorffs α zwischen .92 und .98 erreichen. Grundlage waren zehn schriftliche Reflexionen mit einem Umfang von je 95 bis 121 Textzeilen. Es zeigte sich, dass Feedback zu einem signifikanten Anstieg des Reflexionsniveaus mit Effektstärken zwischen $d = 0.35$ und $d = 0.49$ führt. Jahncke et al. (2018) konnten somit nachweisen, dass Feedback zu schriftlichen Reflexionen über Unterrichtsstunden hohen Einfluss auf die Entwicklung der zweiten Reflexionen hat.

Bei Nitsche (2014) beobachtete die Versuchsgruppe Unterrichtsvideos von Seminarteilnehmern und die Kontrollgruppe Unterrichtsvideos von unbekannten Lehrerinnen und Lehrern. Die VG (Versuchsgruppe) erhielt zusätzlich zur Beobachtung Feedback. Die Ergebnisse der anschließenden schriftlichen Reflexionen zeigten, dass Lehramtsstudierende in der VG ($M = 4.96$; $SD = 0.43$) einzelne unterrichtsrelevante Aspekte genauer beobachteten als die Kontrollgruppe ($M = 4.67$; $SD = 0.44$). Die Ergebnisse waren signifikant mit einem mittleren Effekt ($t_{118} = 3.62$; $d = 0.67$; $p < .001$). Das Ergebnis ist jedoch wegen des niedrigen Wertes für Cronbach Alpha ($\alpha = .52$) vorsichtig zu interpretieren. Weiter konnte gezeigt werden, dass die Teilnehmer der Versuchsgruppe ($M = 4.72$; $SD = 0.60$) die Beobachtungen besser für die Reflexionen nutzen konnten als die Kontrollgruppe ($M = 4.31$; $SD = 0.59$) ($t_{118} = 3.82$; $d = 0.70$; $p < .001$). Auch hier müssen Einschränkungen vorgenommen werden. Der Effekt stammt laut Autor vorwiegend aus der differenzierten Einschätzung zum Item Unterricht. Das Item Lehrerverhalten ist alleine nicht signifikant. Weiter wurde untersucht, ob Beobachtungen einer Videoaufzeichnung einer bekannten Lehrperson einen größeren Effekt haben als Beobachtungen fremder Lehrpersonen. Auch hier konnte die VG ($M = 4.75$; $SD = 0.43$) signifikant bessere Ergebnisse erzielen als die Kontrollgruppe ($M = 3.75$; $SD = 0.67$), mit großem Effekt ($t_{96} = 9.72$; $d = 1.82$; $p < .001$). Die Ergebnisse sind jedoch kritisch zu sehen, erhielt doch die Versuchsgruppe Videos der Seminarteilnehmer und zusätzlich Feedback und die Kontrollgruppe lediglich aufgezeichnete Videos unbekannter Lehrpersonen ohne Feedback. Somit ist nicht genau zu sagen, inwiefern welche Variable Einfluss auf die Ergebnisse

hatte. Ebenso wurde die Reflexion lediglich mit einem Fragebogen untersucht und kann somit nicht alle Facetten des Konstrukts Reflexionskompetenz abdecken. Beispielhaft werden drei Items zitiert:

- a) Ich habe über mein Handeln als Lehrkraft nachgedacht.
- b) Ich habe über die Qualität meiner eigenen Unterrichtsplanung nachgedacht.
- c) Ich habe über meine Eignung als Lehrkraft nachgedacht. (Nitsche, 2014, S. 152)

Die Einschätzung wurde jeweils auf einer sechsstufigen Likert-Skala vorgenommen (1 = „trifft nicht zu“ bis 6 = „trifft voll zu“).

3.4. Messung der Reflexionskompetenz

In Kapitel 3.4.1. werden ausgewählte Modelle zur Messung der Reflexionskompetenz aufgeführt, welche als Grundlage der Entwicklung des Kodiermanuals dieser Studie dienten. Dabei werden die Manuale kurz vorgestellt und Ergebnisse präsentiert. In Kapitel 3.4.2. werden u. a. auch aktuelle Forschungsansätze vorgestellt und so weit möglich Ergebnisse berichtet.

3.4.1. Modelle zur Messung der Reflexionskompetenz

Zeichner & Liston

1985 orientierten sich Zeichner & Liston unter anderem an den drei Kategorien von van Manen (1977) (Grundmodell: drei Stufen, technical rationality (TR), practical action (PA) und critical reflection (CR)) und entwickelten bei der Analyse ihrer Aufnahmen neue Kategorien, wenn bestehende Manuale nicht mehr ausreichten. Sie erarbeiteten vier Stufen des „reflektierenden Lehrindex“ (Tabelle 11). In der ersten Stufe des sachlichen Diskurses (Factual Discourse) beschäftigt man sich lediglich mit dem, was sich primär in der Unterrichtssituation ereignet hat oder damit, was in der Zukunft geschehen wird. Im vernunftbezogenen Diskurs (Prudential Discourse) werden Verbesserungsvorschläge erwogen und das Erreichte evaluiert. Der rechtfertigende Diskurs (Justificatory Discourse) fokussiert auf die angeführten Gründe und wägt Möglichkeiten mit Vor- und Nachteilen ab. In der höchsten Stufe des kritischen Diskurses (Critical Discourse) werden die Gründe untersucht und bewertet, die als Rechtfertigung für pädagogisches Handeln angeführt werden. Ebenso unterzieht man die Werte und Annahmen, die in Form und Inhalt in das Curriculum eingebettet sind, einer Bewertung.

Tabelle 11: Four types of practical discourse (Zeichner & Liston, 1985, S. 162 ff.)

Main level	Lower level	Meaning
Factual Discourse	- Descriptive Discourse - Informational Discourse - Hermeneutic Discourse - Explanatory/Hypothetical Discourse	is concerned with what has occurred in a teaching situation or with what will occur in the future;
Prudential Discourse	- Instruction - Advice/Opinion - Evaluation - Support	revolves around suggestions about what to do or around evaluations of what has been accomplished;
Justificatory Discourse	- Pragmatic Rationale - Intrinsic Rationale - Extrinsic Rationale	focuses on the reasons employed when answering questions of the

		form, “Why do this rather than that?”;
Critical Discourse	- Pragmatic - Intrinsic - Extrinsic - Hidden curriculum	examines and assesses the adequacy of the reasons offered for the justification of pedagogical actions or assesses the values and assumptions embedded in the form and content of curriculum and instructional practices (e.g., the “hidden curriculum”).

Ergebnisse:

In dieser Studie kodierten zwei Personen Tonbandaufnahmen einer Supervisionssitzung. Das Kategoriensystem wurde theoriebasiert (u. a. van Manen, 1977) und aus dem erhobenen Datenmaterial entwickelt. Nach Überprüfung der Interraterreliabilität konnten sie den Hauptanteil der Reflexionen (63,2 %) auf dem niedrigsten Niveau verzeichnen. 24,9 % wurden dem Prudential Discourse zugeordnet, 11,3 % erreichten den Justificatory Discourse und nur 0,6 % der Probanden reflektierten auf dem höchsten Niveau, dem Critical Discourse (Zeichner & Liston, 1985, S. 165).

Van Manen

1991 definierte van Manen Reflexion als eine Aktion, bei der sich der Lernende distanziert, um eine objektivere Sichtweise auf Geschehnisse zu erlangen (van Manen, 1991). Er ordnete Reflexionen in ein kognitiv-hierarchisches Modell, welches von vielen anderen später für ihre Studien angeglichen wurde (Hatton & Smith, 1995) (Tabelle 12).

Tabelle 12: Ebenen der Reflexion von van Manen, zitiert nach Shiel und Jones (2003, S. 6)

Erste Stufe	Tägliches Denken und Handeln nach gesundem Menschenverstand, klare Trennung zwischen Reflexion und Aktion
Zweite Stufe	Spezifischere Reflexion mit Schwerpunkt auf Ereignisse oder Vorfälle
Dritte Stufe	Systematischere Reflexion über die eigene Erfahrung und die der anderen mit dem Ziel, durch Interpretation zu einem Verständnis zu gelangen.
Vierte Stufe	Reflexion über die Art und Weise der Reflexion, Nachdenken über die Natur des Wissens (Meta-Kognition) und die Bedingungen, welche zu Erfahrung führen

Van Manens Werke gelten als die bedeutendsten in der Literatur zur Unterrichtsreflexion (Ballard, 2006; Hatton & Smith, 1995; Shiel & Jones, 2003) und beeinflussten zahlreiche Kodiermanuale und Arbeiten zum Thema Reflexion.

Ergebnisse:

Eine genaue Ausführung der Stufen sowie die zugrunde gelegten Kriterien werden nicht ausgeführt; deshalb ist dieses Modell lediglich als Anhaltspunkt für diese Arbeit zu betrachten (Brendel, 2017).

Hatton und Smith

Eine Vielzahl der Kodiermanuale bezieht sich auf die Ausarbeitungen von Hatton und Smith (1995), welche vier verschiedene Reflexionsstufen unterscheiden. Auf der ersten Stufe, Deskriptives Schreiben (descriptive writing), findet noch keine Reflexion statt. Hier werden lediglich Ereignisse berichtet, auch in Kombination mit passender Literatur, also eine Art Reproduktion der Geschehnisse. Die zweite Ebene, deskriptive Reflexion (descriptive reflection), bezieht sich auf Begründungen, die sich auf die persönlichen Einstellungen bzw. Meinungen beziehen oder auf Literaturrecherche. Auf der dritten Ebene, dialogische Reflexion (dialogic reflection), findet eine Form der Diskussion mit sich selbst statt, eine Suche nach möglichen Ursachen. Die höchste Stufe, kritische Reflexion (critical reflection), definieren die Autoren als Entscheidungen, welche begründet werden können unter Einbezug geschichtlicher, sozialer und/oder politischer oder kultureller Kontexte (Hatton & Smith, 1995). Nach intensiven Aufarbeitungen der bis zu diesem Zeitpunkt veröffentlichten Literatur sehen Hatton und Smith die Notwendigkeit eine genauere Ausführung (Hatton & Smith, 1995, S. 40) zu ergänzen (Tabelle 13). Diese Ebenen werden nicht als hierarchisch verstanden, sondern als Entwicklungsprozess, der idealerweise mit der höchsten Stufe endet (Brendel, 2017; Wischmann, 2015).

Tabelle 13: Types of reflection (Hatton & Smith, 1995, S. 45)

"Reflection type"	Nature of reflection	Possible content
"Reflection-in-action" (Schön, 1983, 1987) addressing IMPACT concerns after some experience in the profes- sion	5. Contextualization of multiple viewpoints drawing on any of the possibilities 1-4 below applied to situations as they are actually taking place	Dealing with on-the-spot professional problems as they arise (thinking can be recalled and then shared with others later)
	4. Critical (social reconstructionist), seeing as problematic, according to ethical criteria, the goals and practices of one's profession	Thinking about the effects upon others of one's actions, taking account of social, political and/or cultural forces (can be shared)
	3. Dialogic (deliberative, cognitive, narrative) weighing competing claims and viewpoints, and then exploring alternative solutions	Hearing one's own voice (alone or with another) exploring alternative ways to solve problems in a professional situation
Reflection-on-action (Schön, 1983; Smith & Lovat, 1990; Smith & Hatton, 1992, 1993) address- ing TASK and IMPACT concerns in the later stages of a preservice program	2. Descriptive (social efficiency, developmental, personalistic), seeking what is seen as 'best possible' practice	Analysing one's performance in the professional role (probably alone), giving reasons for actions taken
	1. Technical (decision making)	Beginning to examine (usually with peers) one's
Technical rationality (Schön, 1983; Shulman,		

1988; Van Mannen [sic], 1977), addressing SELF and TASK concerns early in a program which prepares individuals for entry into a profession	about immediate behaviours or skills), drawn from a given research/theory base, but always interpreted in light of personal worries and previous experience	use of essential skills or generic competencies as often applied in controlled, small scale settings”

Ergebnisse:

Zur Vorbereitung auf ein zehntägiges Praktikum mussten die Probanden ($N = 60$) in der Studie an fünf Microteaching-Sitzungen teilnehmen. Im folgenden Semester hatten sie eine Unterrichtseinheit zu planen und in einem Praktikum durchzuführen und zu evaluieren. Dabei konnten sie ihre Vorhaben mit einem „critical friend“ diskutieren. In einer Hausarbeit wurden die gehaltenen Stunden reflektiert. Fokus wurde auf die Planung und tatsächlich erfolgte Durchführung gelegt. Bezogen auf die Selbstevaluationen, die Videodokumentationen, die Interviews und die Vignettentests, welche alle auf Reflexionskompetenz untersucht wurden, konnten die besten Ergebnisse in den schriftlichen Hausarbeiten gefunden werden. Lediglich eine Hausarbeit enthielt keine reflexiven Anteile. Das Maximum an reflexiven Kodierungen betrug 52, das Minimum lag bei 2. Der Großteil, ca. 60 - 70 % aller Arbeiten, wies die zweite Stufe (descriptive reflection) auf und dies vorwiegend zu Beginn der Arbeiten. Dialogische Ansätze, also Stufe drei, konnten oft auf die Unterstützung des „critical friends“ zurückgeführt werden sowie auf erfolgreiche Diskussionen von Videoausschnitten, und fanden sich am Ende der Hausarbeiten in Form von Alternativvorschlägen. Nur acht Studierende (aus zwei Kohorten) zeigten Ansätze kritischer Reflexion (höchste Stufe). Die Autoren merkten an, dass die Studierenden einen hohen Anteil (fast 50 %) an Perspektivenübernahme erreichten.

Bain et al.

Bain et al. (1999) unterschieden erstmals zwischen Reflexionstiefe und Reflexionsbreite und kombinierten diverse Modelle, um eine fünfstufige Reflexionsskala und zusätzlich vier Kategorien nach inhaltlichen Fokussen von Lerntagebüchern zu erhalten (Tabelle 14).

Tabelle 14: Fünfstufige Reflexionsskala nach Bain et al. (1999, S. 60)

Fünfstufige Reflexionsskala	
Level 1 (Berichterstattung)	Der/die Studierende beschreibt, berichtet oder erzählt mit minimaler Transformation, ohne zusätzliche Beobachtungen oder Einsichten
Level 2 (Antworten)	Der/die Studierende verwendet die gelernte Theorie in irgendeiner Weise, aber mit wenig Transformation oder Konzeptualisierung Der/die Studierende macht eine Beobachtung oder fällt ein Urteil, ohne weitere Schlussfolgerungen zu ziehen oder die Gründe für das Urteil im Einzelnen darzulegen Der/die Studierende stellt eine "rhetorische" Frage, ohne zu versuchen, sie zu beantworten oder Alternativen zu erwägen

	Der/die Studierende berichtet von einem Gefühl wie Erleichterung, Angst, Glück usw.
Level 3 (Bezug nehmen)	Der/die Studierende identifiziert Aspekte, die persönliche Bedeutung haben oder die sich mit früheren oder aktuellen Erfahrungen decken Der/die Studierende sucht ein oberflächliches Verständnis von Beziehungen Der/die Studierende identifiziert etwas, in dem er gut ist, etwas, das er verbessern muss, einen Fehler, den er gemacht hat, oder einen Bereich, in dem er aus seiner praktischen Erfahrung gelernt hat. Der/die Studierende gibt eine oberflächliche Erklärung des Grundes, warum etwas passiert ist, oder er/sie identifiziert etwas, was sie/er tun oder ändern will oder muss
Level 4 (Begründen)	Der/die Studierende integriert die Daten in eine angemessene Beziehung, z.B. mit theoretischen Konzepten, persönlichen Erfahrungen, die ein hohes Maß an Transformation und Konzeptualisierung beinhalten Der/die Studierende sucht ein tiefes Verständnis dafür, warum etwas passiert ist Der/die Studierende erforscht oder analysiert ein Konzept, ein Ereignis oder eine Erfahrung, stellt Fragen und sucht nach Antworten, erwägt Alternativen, spekuliert oder stellt Hypothesen auf, warum etwas geschieht Der/die Studierende versucht, sein/ihr eigenes oder das Verhalten oder die Gefühle anderer zu erklären, indem er/sie seine/ihre eigenen Einsichten, Schlussfolgerungen, Erfahrungen oder sein/ihr früheres Lernen mit einer gewissen Tiefe des Verständnisses verwendet Der/die Studierende setzt sich eingehend mit der Beziehung zwischen Theorie und Praxis auseinander
Level 5 (Rekonstruktion)	Der/die Studierende zeigt ein hohes Maß an abstraktem Denken, um das Gelernte zu verallgemeinern und/oder anzuwenden Der/die Studierende zieht eine originelle Schlussfolgerung aus seinen/ihren Überlegungen, verallgemeinert aus seinen/ihren Erfahrungen, extrahiert allgemeine Prinzipien, formuliert eine persönliche Theorie des Lehrens oder nimmt zu einem Thema Stellung. Der/die Studierende extrahiert und verinnerlicht die persönliche Bedeutung seines/ihrer Lernens und/oder plant sein/ihr eigenes weiteres Lernen auf der Grundlage seiner/ihrer Reflexionen

Die Reflexionsbreite bei Bain (Tabelle 15) ist stark auf pädagogische Handlungen und Einschätzungen der Lehrkraft ausgerichtet (Brendel, 2017):

Tabelle 15: Reflexionsbreite nach Bain et al. (1999, S. 59)

Kategorien, die den Fokus der Reflexion beschreiben	
Schwerpunkt Lehre	Allgemeine Unterrichtsfragen Spezifische Lektionen oder spezifische Vorfälle Verhaltensmanagement Sonstiges (Unterrichtsmanagement, Inhalt, Hausaufgaben)
Fokus auf das Selbst	Eigene Fähigkeiten/Wert als Lehrer Eigener Lehransatz Andere (Gefühle, Identität, Lernen)

Fokus auf berufliche Fragen	Beziehung/Diskussion mit dem betreuenden Lehrer Lehren als Beruf Institutionelle Fragen Professionelle Vorbereitung Soziale/ethische Fragen Aktivitäten außerhalb der Klasse
Fokus auf SchülerInnen und Schüler oder Klasse	Verhalten der Klasse Schüler-/Klassenmerkmale Beziehung zu bestimmten Schülerinnen und Schülern Unterrichten bestimmter Schülerinnen und Schüler

Ergebnisse:

Dieses Modell zur Auswertung von Lerntagbüchern wurde bereits in vielen weiteren Studien verwendet und hat sich als empirisch belastbar und brauchbar zur Bestimmung der Reflexionskompetenz erwiesen (Brendel, 2017; Carrington & Selva, 2010; Hsieh, Jang, Hwang & Chen, 2011; Hung, Yang, Fang, Hwang & Chen, 2014). In den aufgeführten Studien wurde das Kodiermanual nur geringfügig adaptiert, was als Hinweis für die Qualität dieses Modells gewertet werden kann.

Kember

Das Modell von Kember et al. (2000) basiert auf dem linearen Stufenmodell von Mezirow und unterscheidet klar zwischen inhaltlicher und prozessorientierter Reflexion. Zusätzlich werden sowohl verschiedene Qualitäten als auch eine hierarchische Stufung vollzogen. Die sieben Stufen von Mezirow werden in nicht-reflexive und reflexive Stufen unterteilt, die sich wiederum in höhere und niedrigere Unterstufen aufteilen lassen (Brendel, 2017). Diese Einteilung hat man jedoch überarbeitet, da sie von den Autoren als zu fein erachtet wurde. Es fand eine Reduzierung auf folgende vier Stufen statt (Tabelle 16):

Tabelle 16: Stufenmodell nach Kember (Kember et al., 2000, S. 383 ff.)

Ge- wohnheitsmäßige Handlung	Gewohnheitshandlung ist das, was man vorher gelernt hat und durch häufigen Gebrauch zu einer Tätigkeit wird, die automatisch oder mit wenig bewussten Gedanken ausgeführt wird. Wenn eine bestimmte Art von Problem viele Male erlebt wurde, wird die Art, mit ähnlichen Fällen umzugehen, zur Routine. Schön (1983) nannte diese Art von Verhalten "knowing-in action".
Verständnis	Blooms (1979) definiert das Verstehen als "Verstehen, ohne sich auf andere Situationen zu beziehen". Eine akademische Art des Lernens, bei der der Student zu einem Verständnis eines Konzepts gelangen kann, ohne über die Bedeutung in persönlichen oder praktischen Situationen nachzudenken.
Reflexion	Die letztgenannte Definition ist für die Berufspraxis von besonderer Bedeutung, da sie die Erfahrung als Prüfstein für die Reflexion betrachtet. Diese Definition bewegt sich mehr in Richtung des Rahmens von Schön (1983) und des reflektierenden Praktikers im Kontext der Berufspraxis.
Kritische Reflexion	Mezirow verwendet den Begriff der Prämissenreflexion. "Prämissenreflexion bedeutet, dass wir uns bewusst werden, warum wir

	wahrnehmen, denken, fühlen oder handeln, wie wir es tun" (Mezirow, 1990, S. 108). (...) "Sie muss eine Lücke beinhalten, in der ein Problem neu definiert wird, so dass die Aktion neu ausgerichtet werden kann" (Mezirow, 1990, S. 110).

Ergebnisse:

Zur Förderung der Reflexionskompetenz wurden in dieser Studie Lerntagebücher verwendet. Das Manual erreichte eine Interraterreliabilität von $\alpha = 0.65$. Bei Unstimmigkeiten verwendeten die Autoren eine argumentative Validierung. Hierbei konnten unterschiedliche Kodierungen auf unterschiedliche Interpretationen der Tagebucheinträge zurückgeführt werden.

Zimmermann und Welzel

Zimmermann und Welzel (2008) entwickelten die vierstufigen Modelle von Zeichner und Liston (1985) und von Hatton und Smith (1995) induktiv um eine Stufe weiter und beschrieben ihre Kompetenzstufen der Reflexivität wie folgt (Tabelle 17):

Tabelle 17: Modell nach Zimmermann und Welzel (2008)

Reflexionstiefe	Merkmale
• Stufe 1 Sachbezogene Beschreibung	• Bezogen auf konkretes pädagogisches Angebot • Einfache Wiedergabe von Handlungsabläufen • Beschreibung von Situationen ohne zusätzliche Kommentare, ohne eigene Gedanken • Erste Distanzierung vom konkreten Tun ohne Begründung, keine explizite Reflexion
• Stufe 2 Handlungsbezogene Begründung	• Beschreibung von Sachverhalten • Sinngebungen mit Begründungen oder Handlungsalternativen werden aufgeführt. • Innere Distanzierung vom bloßen Handeln, da Formulierung der eigenen Perspektive vorhanden
• Stufe 3 Analytische Abstraktion	• Gemachte Erfahrungen werden unter Berücksichtigung mehrerer Perspektiven (z. B. aus der Sicht von Erzieherin und Kind, Vergangenheit und Gegenwart – in Bezug auf eigenes pädagogisches Handeln) analysiert. • Elemente und Strukturen werden auf einer allgemeineren Ebene erkannt, identifiziert und zueinander in Beziehung gesetzt. • Aus dieser Analyse heraus werden Konsequenzen im Hinblick auf zukünftiges Verhalten, zukünftige Aktivitäten gezogen. • Ein neues erweitertes Verständnis wird entwickelt.
• Stufe 4 Kritischer Diskurs	• Die Erkenntnisse werden im Hinblick auf übergeordnete Werte, Prinzipien und/oder auf relevante Theorien hin bewertet.

	• Es erfolgt eine kritische Auseinandersetzung damit. • Es kommt zu einer Zunahme der Fähigkeit zum Erkennen eigener Anteile und Deutungen in Interaktionen.
• Stufe 5 Professionalisierung	• Erkennen des eigenen Urteils-Handlungs-Habitus und Durchdenken, Planen von zielorientierten Verbesserungsstrategien • Konkrete Verbesserungsmaßnahmen (am eigenen Tun, am professionellen Umfeld) werden initiiert. • Festgefahrene Rollenverständnisse werden thematisiert und konstruktiv in Frage gestellt, andere Handlungsmöglichkeiten diesbezüglich ausgelotet und ausprobiert.

Ergebnisse:

Die Autoren untersuchten Interviews ($N = 16$) auf Reflexionstiefe und inhaltliche Komponenten. Hierbei konnte eine höhere Stufe nur erreicht werden, wenn alle Punkte der vorhergehenden Stufe erfüllt waren. Die Übereinstimmung des Manuals war hierbei mit $\kappa = .94$ äußerst zufriedenstellend. Eine Wiederverwendung dieses Stufenmodells bei anderen Professionalisierungsprogrammen scheint nach Abels (2011) aufgrund der Fünfstufigkeit und der Komplexität schwer umsetzbar und wird deswegen nur selten praktiziert.

Leonhard & Rihm

Das Modell von Leonhard et al. (2011) (Tabelle 18) beschreibt zwei miteinander verschränkte Analyseperspektiven: zum einen die Reflexionsbreite und zum anderen die Reflexionstiefe. Ersteres beschreibt die Aspekte, welche eine Lehrperson einbezieht. Bei Zweitem orientieren sich Leonhard und Rihm (2011) an dem Modell von Hatton und Smith (1995). Als Gründe nennen sie die Entwicklung des Modells auf Basis der Analyse schriftlicher Reflexionen, die Entstehung im Kontext der Lehrerbildung und die inkludierte Entwicklungsperspektive, welche die Stufen als Formen zunehmender Komplexität ansieht (Gehrmann et al., 2010) (Tabelle 18).

Tabelle 18: Stufenmodell zur Einschätzung der Reflexionstiefe (Leonhard & Rihm, 2011, S. 257)

Stufe	Bezug zu Hatton/Smith	Situationswahrnehmung	Darstellung der Handlungsoptionen
0	descriptive writing	Schlichte Beschreibung der Situation	Benennung einer Handlung
1	descriptive reflection	Situationsbeschreibung mit Entwicklung einer eigenen Annahme zur Situation (Annahme: implizit)	Beschreibung einer Handlung mit möglichen Alternativen

2		Situationsdeutung, in der zwei Annahmen einfach miteinander verknüpft werden	Beschreibung und Begründung einer Handlung auf der Basis impliziter Annahmen und Mentalitäten bzw. subjektiver Handlungsnormen
3		Situationsdeutung durch Verknüpfung von zwei und mehr subjektiven Hypothesen, die begründet miteinander verknüpft werden (subj. Hypothese: explizite Annahme)	Beschreibung und Begründung einer Handlung auf der Basis expliziter Annahmen, aktive Distanzierung wird deutlich
4	dialogic reflection	Einbezug des Kontextes für die komplexe Einschätzung der Situation	Explizierte subjektive Theorie
5		Situationsanalyse unter Einnahme einer weiteren Perspektive (Perspektivenübernahme)	Explizierte subjektive Theorie sowie Einbezug weiterer Perspektiven
6		Einbezug theoretisch-wissenschaftlicher Kategorien oder Ergebnisse zur Analyse der Situation unter der Perspektive der Bedingungsverfügung	Explizierte subjektive Theorie sowie Einbezug weiterer theoretisch-wissenschaftlicher Perspektiven
7	critical reflection	Dialektische Auseinandersetzung mit der eigenen Situationswahrnehmung, theoretischen Bezügen und institutionellen Erwartungen	Dialektische Auseinandersetzung der eigenen Position mit weiteren, auch wissenschaftlichen Theorien

Ergebnisse:

Dieses Evaluationsprojekt untersucht die Auswirkung einer Begleitveranstaltung zu schulpraktischen Studien auf die Reflexionskompetenz. Das an Studierendendaten entwickelte Reflexionstiefenmodell konnte in der Breite der Daten und über die drei verschiedenen Reflexionsvignetten nicht empirisch tragfähig abgebildet werden. Es zeigten sich sowohl bei der Untersuchungsgruppe als auch bei der Vergleichsgruppe ein „Kompetenzverlust“, wobei dieser in der Untersuchungsgruppe signifikant niedriger ausfällt als in der Vergleichsgruppe ($t_{94} = 2.72, p = .007$). Auch die Effektstärke ist mit $d = 0.90$ bei der Vergleichsgruppe deutlich größer als bei der Untersuchungsgruppe mit $d = 0.15$. Die Ursachen dafür können nach der Analyse der Untersuchungsebenen vielfältig sein und zeigen keine einheitliche Tendenz. Die Zeichenzahl als Indikator des Umfangs korreliert hierbei zu beiden Zeitpunkten mit beiden Dimensionen, also Reflexionsbreite und Reflexionstiefe, wesentlich bis hoch. Die Autoren ziehen hierbei Konsequenzen, indem sie die Gestaltung der Reflexionsvignetten optimieren sowie zusätzlich die Motivation der Teilnehmer erheben, um die Wahl des Untersuchungszeitpunktes als zentrale Einflussgröße zu belegen.

Bräuer und Schindler

Bräuer und Schindler (2011) werteten ihre Portfolios anhand der folgenden Ebenen der Reflexion aus (Tabelle 19):

Tabelle 19: Die Ebenen der reflexiven Praxis (Bräuer, 2016, S. 28)

Ebenen der Reflexion	4	Planen	... von Handlungsalternativen
	3	Beurteilen	... auf Basis (an)erkannter Kriterien
		Bewerten	...im Vergleich mit Erwartungen bzw. anderen Leistungen
	2	Interpretieren	...mit Blick auf die Konsequenzen aus der eigenen Handlung
		Analysieren	...mit Bezug zur eigenen Leistung
	1	Dokumentieren	...mit Bezug zur Gesamthandlung
		Beschreiben	Tätigkeitsbeschreibung

Ergebnisse:

Die Reflexion in Portfolios verbleibt zumeist auf der Ebene der Dokumentation der Handlungen (Keller & Bräuer, 2013). Bräuer (2016, S. 28 f.) nimmt an, „dass diese Schwierigkeiten beim Reflektieren durch weniger erfahrene Schreibende mit deren Stand der allgemeinen mentalen Entwicklung zusammenhängen und in diesem Kontext auch den Stand der sprachlichen Entwicklung bzw. das Fehlen rhetorischer Mittel zum differenzierten reflexiven Gebrauch von Sprache demonstrieren.“ Um das erreichte Niveau zu verbessern, sollten spezielle Aufgaben gestellt werden, welche die komplexe Herausforderung des Reflektierens erleichtern oder in Zwischenschritte zerlegen. Unterstützt wurde der Übergang von einer Reflexionsebene auf die nächste durch Peer-Feedback bzw. Schreibberatung. Die Auswertung der Feedback-Protokolle, Textrevisionen und Interviews ergab, dass das Sprachmaterial für die nächsthöhere Reflexionsebene zwar vorhanden war, es aber nicht zum Einsatz kam, da der kommunikative Bedarf nicht erkannt wurde. Deutlich wurde dies den meisten Studierenden erst auf entsprechende Nachfrage hin (z.B. durch Peer-Feedback).

Wyss

Corinne Wyss (2013) führte in ihrer Studie unter anderem Stimulated Recall Interviews durch und wertete diese nach folgendem Schema aus:

Teil A des Analyseinstrumentes umfasste allgemeine Auswertungen wie die Anzahl der Wörter, die Anzahl der Bedeutungseinheiten der Aussagen, den Fokus auf sachliche Inhalte, den Fokus auf Lehr-Lern-Methoden und -Strategien und allgemeine Aussagen. So wurde ein allgemeiner Überblick über die Beschaffenheit der Daten gewonnen. Wyss betont hierbei, dass besonders durch die genannten Inhalte, Methoden, Strategien und allgemeinen Aussagen bereits Tendenzen der Gestaltung von Reflexionsprozessen zu erkennen waren.

Teil B teilte sich in drei Unterthemen:

B1 Beurteilung der Qualität des Unterrichts

Hierunter fallen zum Beispiel Punkte wie Qualität der Organisation, Diagnostik, Klassenführung, Motivierungsfähigkeit, Sachkompetenz, Strukturierung und Differenzierung. In dieser Studie fällt diese Art der Auswertung unter die Begrifflichkeit Reflexionsbreite. Ebenso wird bei Wyss wie auch in dieser Studie eine Unterscheidung vorgenommen, ob die Lehrperson einen Punkt als positiv oder negativ wahrnimmt.

B2 Fokus der Lehrpersonen bei der Reflexion

Wie auch bei Harrington et al. (1996) wurde untersucht, auf welcher Person der Fokus der Aussage liegt. Unterschieden wurden die Lehrperson selbst, die Schülerinnen und Schüler und Drittpersonen wie beispielsweise Eltern oder andere Lehrpersonen. In dieser Studie wurde dieser Punkt als Perspektivenübernahme bezeichnet und unterscheidet die Lehrerperspektive, die Schülerperspektive, die Fachperspektive, die Theorieperspektive und die Sozial-Politische Perspektive.

B3 Benennung von Handlungsalternativen

In Anlehnung an Fischer, Rieck und Lobemeier (2008) untersuchte man die Reflexionsbeiträge auf alternative Handlungsmöglichkeiten.

B4 Qualität des Reflexionsprozesses (Tabelle 20)

Es wurde eine fünfstufige Likert-Skala verwendet. Hierbei stellt die Stufe 1 das niedrigste Level einer Reflexion dar und folglich Stufe 5 das höchste.

Tabelle 20: Kodiermanual (Wyss, 2013, S. 128)

Deskription	Reine Beschreibung. Die Ausführungen der Lehrperson sind rein deskriptiv und beinhalten keine Erklärungen oder Begründungen.
Persönliche Erläuterungen	Erklärungen, Beschreibungen mit Bezug auf persönliche Erfahrungen. Die Ausführungen der Lehrperson beinhalten Erklärungen, die auf persönlichen Erfahrungen beruhen.
Theoretische Erläuterungen	Erklärungen, Beschreibungen mit Bezug auf theoretisches Vorwissen.
Kritische Erläuterungen	Umfangreichere Erklärungen mit Bezug auf Theorie, Kontext und daraus zu schließenden Folgerungen.
Ethisch moralische Erläuterungen	Erklärung mit Berücksichtigung von ethischen, moralischen, politischen, gesellschaftlichen Gesichtspunkten.

Ergebnisse (Tabelle 21):*Tabelle 21: Ergebnisse der prozentualen InterCoderübereinstimmung und der InterCoder-reliabilität nach Fleiss' Kappa*

Teil	Prozentuale InterCoderübereinstimmung	Fleiss' Kappa
B1	83 %	.74
B2	98 %	.96
B3	94 %	.85
B4	90 %	.76

Beim Vergleich zwischen jungen (zwei Messzeitpunkte; erstes Berufsjahr) und erfahrenen Lehrpersonen (Kontrollgruppe; mindestens sechs Jahre Berufserfahrung), welche lediglich als Referenzgruppe dienten, war kein generalisierbarer Unterschied bei der Anzahl von Bedeutungseinheiten zu finden. Tendenziell konnten aber umfangreichere und einheitlichere Aussagen bei den Junglehrpersonen gemessen werden. Bei den erfahrenen Lehrpersonen wurde ein etwas höherer Anteil an sachinhaltlichen Aussagen gemacht, es fanden sich jedoch keine signifikanten Unterschiede der beiden Gruppen.

Die Beurteilung der Qualität des Unterrichts bezieht sich bei allen befragten Lehrpersonen schwerpunktmäßig auf die Punkte „Qualität der Organisation“ (20 Prozent), „Motivierungsfähigkeit“ (10 Prozent), und „Klassenführung“ (10 Prozent). Im Verlauf der Studie konnte bei den Junglehrern von Messzeitpunkt 1 zu Messzeitpunkt 2 eine Zunahme für die „Motivierungsfähigkeit“ und eine Abnahme bei der „Klassenführung“ verzeichnet werden. Junglehrer erwähnten zusätzlich mit 10 Prozent den „Umgang mit Unterrichtszeit“, was bei den erfahrenen Lehrpersonen mit nur 4 Prozent kaum eine Rolle zu spielen scheint. Wyss gibt jedoch zu bedenken, dass diese Ergebnisse nicht statistisch abgesichert sind. 20 Prozent der Aussagen, sowohl bei jungen als auch bei erfahrenen Lehrpersonen, ließen sich keinem vordefinierten Qualitätskriterium zuordnen und wurden in der Kategorie „Anders“ verortet. Infolge dessen wurden häufig allgemeine Kommentare genannt, die sich nicht auf spezifische Qualitätskriterien beziehen, wie sie in dieser Studie von Wyss definiert werden.

Am Anfang des Berufsjahres konnte der höchste Anteil an negativen Beurteilungen bei den jungen Lehrpersonen gemessen werden. Diese ändern sich von Messzeitpunkt 1 mit 30 Prozent auf Messzeitpunkt 2 mit 26 Prozent. Allgemein ist der größte Anteil der gemachten Beurteilungen neutral/deskriptiv (Junglehrer: 49 Prozent bei Messzeitpunkt (MZP) 1 und 51 Prozent bei Messzeitpunkt (MZP) 2; erfahrene Lehrpersonen: 54 Prozent). Jedoch sind auch die hier erhaltenen Ergebnisse statistisch nicht bedeutsam.

Zum Fokus der Aussage: Durchschnittlich waren 80 Prozent der Aussagen auf die eigene Person bezogen und etwa 20 Prozent bezogen sich auf die Schülerinnen und Schüler, wobei sich bei den Junglehrern hier ein kleiner Zuwachs von MZP 1 (ca. 18 Prozent) auf MZP 2 (ca. 22 Prozent) ergab.

Handlungsalternativen wurden von den befragten Lehrpersonen selten benannt. Das Maximum der genannten Alternativen ergibt sich bei den Junglehrern bei MZP 1 mit 22 Prozent. Dies nimmt zum MZP 2 auf 18 Prozent ab. Bei der Kontrollgruppe konnten nur

15 Prozent gemessen werden. Auch hier sind die berichteten Ergebnisse nicht statistisch abgesichert.

Die Studierenden konnten bei der Qualität des Reflexionsprozesses nur die Stufe 1 und Stufe 2, also rein deskriptiv oder persönliche Erläuterungen, erreichen. Höhere Stufen wurden in dieser Studie nicht kodiert. Den größten Anteil macht hier die Stufe 2 aus. Die Junglehrer erreichten dabei im ersten MZP einen Wert von 63 Prozent, welcher bis zum MZP 2 auf 74 Prozent anstieg. Den größten Anteil erreichten jedoch die erfahrenen Lehrkräfte mit 79 Prozent. Auch diese Ergebnisse sind statistisch nicht von Bedeutung. Weitere Ergebnisse dieser Studie werden im nächsten Kapitel vorgestellt.

Windt

Windt, Hasenkamp, Rau, Lenske und Rumann (2016) untersuchten 25 Lehramtsanwärterinnen und Lehramtsanwärter über drei Messzeitpunkte. Nach einer gezeigten Unterrichtsstunde hatten die Probanden zehn Minuten Zeit über die Stunde nachzudenken und gaben dann eine ca. zehnminütige mündliche Reflexion ab. Diese wurde audiografiert, transkribiert und anhand des folgenden Kategoriensystems analysiert (Tabelle 22):

Tabelle 22: Kategoriensystem (Windt & Lenske, 2015a, S. 214)

Kategorie	Kodierung
Inhalt	Facetten aus den Kategoriensystemen für Planung und Durchführung
Veridikalität	5-stufige Bewertung durch FachleiterInnen und MentorInnen
Niveaustufen	1. Beschreibung 2. Bewertung 3. Begründung 4. Alternativen 5. Konsequenzen
Objektivität	Dichotom für jede Facette, dann Durchschnitt über alle Facetten
Strukturierung	Dichotom für jede Facette, dann Durchschnitt über alle Facetten
Umfang	• Wortanzahl • Anzahl an genannten Facetten • Anzahl an Fachbegriffen

Ergebnisse (Abbildung 22):

Das Kodiermanual erreichte sehr gute Werte mit $\kappa = .96$. Die Autoren fanden große Unterschiede bei den angehenden Lehrpersonen. So waren manche fähig, eine hohe Reflexionstiefe zu erreichen, andere beschrieben lediglich bestimmte Situationen. Das erreichte Level der Reflexionen am Anfang der Studie korrelierte nicht mit der Anzahl an genannten Aspekten, spricht der Reflexionsbreite.

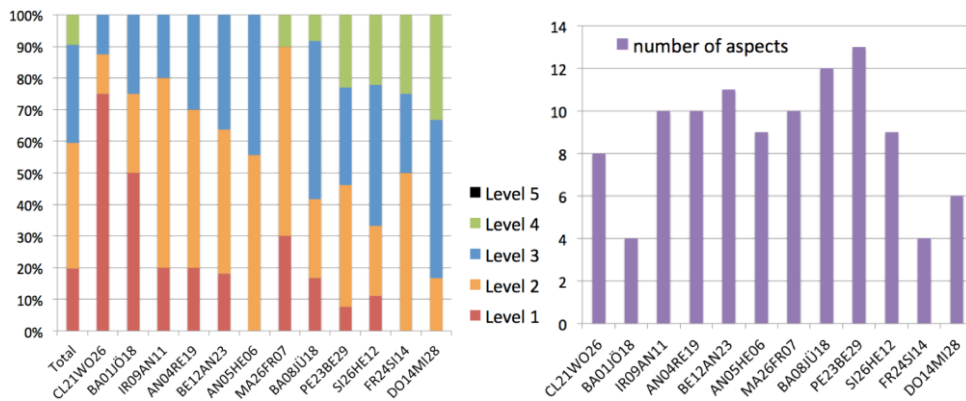


Abbildung 22: Ausgewählte Ergebnisse der Studie (Windt et al., 2016, S. 2193)

Die Autoren verwiesen allerdings auf die kleine Stichprobenzahl, welche keine wirklich aussagekräftigen Schlussfolgerungen zulässt.

Zusammenfassung (Tabelle 23):

Tabelle 23: Zusammenfassung und Bezug auf das selbst erarbeitete Kodiermanual

Autoren	Literaturbezug
Zeichner und Liston (1985)	Bewertung von Verbesserungsvorschlägen, Begründungen, Abwägen, Theoretische Begründungen, Bezüge zum Lehrplan
van Manen (1991)	- Modell über vier Stufen - Der Lernende distanziert sich, um eine objektivere Sichtweise auf Geschehnisse zu erlangen.
King und Kitchenner (1994)	In der höchsten Stufe des Reflective Thinking werden Perspektiven als Kontext zur Erlangung von Wissen betrachtet.
Hatton und Smith (1995)	- Reine Beschreibung als unterste Stufe - Begründungen, eigene Meinung, Literaturrecherche - Anführen möglicher Ursachen mit Abwägen - Begründung unter Einbezug weitreichender Überlegungen - Zielformulierungen
Bain et al. (1999)	- Unterscheidung der Reflexionstiefe und Reflexionsbreite - Die Reflexionsbreite beinhaltet den Fokus auf die Lehrperson selbst, auf das Unterrichten, auf berufliche Fragen und auf die Schülerinnen und Schüler
Kember et al. (2000)	- Beleg, dass eine größere Unterteilung (in sieben Stufen) nicht zielführend ist; Modell bleibt bei vier Stufen - Unterteilung in „nicht reflexiv“ und „reflexiv“
Zimmermann und Welzel (2008)	Stufe 1: Einfache Beschreibungen ohne Reflexion Stufe 2: Begründungen und Handlungsalternativen Stufe 3: Perspektivenübernahme; Konsequenzen werden gezogen;

	Verbesserungsvorschläge gebracht Stufe 4: Theoriebezug bzw. Literaturbezug Stufe 5: Verbesserungsstrategien; Handlungsalternativen → Perspektivenübernahme ist zur Erlangung eines höheren Niveaus nötig.
Leonhard und Rihm (2011)	- Unterscheidung von Reflexionsbreite und Reflexionstiefe - Basis bilden die vier Stufen von Hatton und Smith. - Inhalt der schriftlichen Reflexionen: Schlichte Beschreibung; Entwicklung einer eigenen Annahme; Beschreibung und Begründung einer Handlung; explizierte subjektive Theorie; Einbezug weiterer Perspektiven; Einbezug weiterer theoretisch-wissenschaftlicher Perspektiven; wissenschaftliche Theorien
Bräuer und Schindler (2011)	- Der Einsatz eines Portfolios wird verworfen und schriftliche Reflexionen beschlossen. - Gefördert werden die Reflexionen einer Interventionsgruppe durch Feedback eines ausgebildeten Mitarbeiters. - Vier Phasen: Beschreiben, Bewerten, Beurteilen und Planen von Handlungsalternativen
Wyss (2013)	- Untersuchung der Reflexionsbreite mit induktiv und deduktiv entwickelten Kategorien, um alle genannten Bereiche abdecken zu können. - Untersuchung der Reflexionsbreite durch ein Punktesystem, verworfen wird ein hierarchisches Modell der Einschätzung der Reflexionstiefe. - Handlungsalternativen werden nicht höher bewertet. - Der Literaturbezug als wesentlicher Punkt einer Reflexion wird aufgenommen.
Windt et al. (2016)	Untersuchung der folgenden Punkte: Beschreibung, Bewertung, Begründung, Alternativen, Konsequenzen

3.4.2. Wirksamkeit von Reflexionsprozessen

Bisherige Ergebnisse der veröffentlichten empirischen Studien, welche die Effekte der Programme zu reflexivem Denken messen, sind enttäuschend. So berichten Chandler, Robinson und Noyes (1990), dass Reflexion nicht durch Training verbessert werden kann. 1991 verzeichnen die Autoren in ihrer Vergleichsstudie, dass die erfolgreichsten Studierenden jene waren, welche innovativ ihren Unterricht planten und willens waren, Dinge auszuprobieren. Ihre Maßnahme, eine Vergleichsgruppe kritisch reflektieren zu lassen, konnte nicht überzeugen. Sie führten dies auf das frühe Stadium zurück, in welchem Studierende in der Lehrerausbildung sind, oder auf die fehlende bzw. falsche Definition des Begriffs „Reflexion“.

El-Dib (2007) untersuchte die Reflexionsebenen in der Aktionsforschung (Levels of reflection in action research). Ausgewertet wurden 163 Studierende im dritten Jahr der Ausbildung und 159 Studierende des vierten Jahres. Es begann mit einer Einführung in die Aktionsforschung (Kapitel 3.1.2), darauf folgte die Aufgabenstellung, selbst im nächsten

Unterrichtspraktikum eine kleine Untersuchung im Sinne der Aktionsforschung durchzuführen. Hierbei durften die Probanden ihr Thema frei wählen und reflektierten ohne jegliche Inputphasen oder andere Unterstützungen. Die Portfolios wurden im Anschluss auf die Reflexionstiefe untersucht. Integriert wurden die drei Phasen der Aktionsforschung planen, handeln und überprüfen, welche wiederum auf vier Niveaustufen (niedrig, niedrig mittel, hoch mittel, hoch) reflektiert werden konnten. Die Ergebnisse waren ernüchternd (Tabelle 24):

Tabelle 24: Ergebnisse der Studie von El-Dib (2007)

Stages of action research	Levels of reflective thinking			
	low in %	low medium in %	high medium in %	high in %
Planning	67	19	12	2
Plan of action	30	43	21	6
Acting	39	48	11	2
Reviewing	30	29	29	2
Overall	30	65	4	1

Da die Studierenden am Ende ihrer Ausbildungszeit waren, wurde anfangs beschlossen, dass mindestens 50 Prozent oder mehr der Studierenden die Level higher medium oder high erreichen sollten. Die Leiter der Studie mussten jedoch feststellen: „The finding that more than 50 Prozent of the prospective teachers participating in this study fell at either the *low* or *low-medium* level of reflection in all the stages of action research is rather alarming“ (El-Dib, 2007, S. 32). In der Diskussion des Beitrages schlugen sie Fördermaßnahmen vor wie Tagebuch schreiben, Reflexionsteams und Rollenspiele. „Scaffolding and mentoring are required in order to develop prospective teachers’ reflective abilities given the rather complicated and difficult nature of this process“ (El-Dib, 2007, S. 33).

Das Kodiermanual von Wyss (2013) und die darin erzielten Ergebnisse wurden im letzten Kapitel bereits erläutert. Weiter wurden mittels Videografie die Unterrichtsstunden kodiert und geratet und so analysiert. Zudem setzte man Vignetten mit Problemsituationen ein, für welche es galt Lösungen zu finden, die wiederum ausgewertet wurden. Die Schülerinnen und Schüler bekamen Fragebogen, die Lehrpersonen Reflexionsbögen und der Fragebogen NEO-FFI, in welchem Persönlichkeitsmerkmale wie „Offenheit für Erfahrungen“ untersucht werden, wurde zusätzlich ausgefüllt. Zentrales Ergebnis dieser Studie ist, dass die Reflexionsfähigkeit der befragten Lehrpersonen als wenig entwickelt beurteilt wird. Hierbei unterscheidet sich das Reflexionsverhalten von Junglehrern am Anfang und am Ende des ersten Berufsjahres statistisch nicht von erfahrenen Lehrpersonen. Die Selbstbeurteilung des Unterrichts von den Lehrpersonen weicht von der Fremdeinschätzung des Unterrichts durch die Schülerinnen und Schülern sowie Expertinnen und Experten ab. Auch konnte die Analyse der Daten zeigen, dass in verschiedenen Bereichen Zusammenhänge zwischen Reflexionsverhalten und Persönlichkeitsmerkmalen bestehen. Allerdings weichen die Ergebnisse hierbei von den Erwartungen ab. Zu Deweys (1933) drei Voraussetzungen für Reflexionskompetenz (siehe Kapitel 3.2.2) lassen sich keine Belege finden. Allerdings konnten Zusammenhänge zwischen der Skala „Offenheit für Erfahrungen“ und der Fremdbeurteilung der Reflexion im Stimulated Recall Interview ausgemacht werden. Ebenso wurde ein negativer Zusammenhang für die Skala „Gewissenhaftigkeit“ (Anstrengungsbereitschaft, Ehrgeiz, Disziplin) und den Stimulated Recall

Interviews herausgearbeitet. Probanden mit hohen Ausprägungen auf der Gewissenhaftigkeitsskala äußerten sich im Interview quantitativ weniger ausführlich (das Alpha-Fehlerniveau lag bei $p = .056$, also statistisch knapp nicht signifikant). Als nachvollziehbar werden die Zusammenhänge der Skala „Gewissenhaftigkeit“ und der Beurteilungen des Verhaltens von Personen bei Junglehrern bewertet.

Eines der immer wieder auftretenden Probleme bei der Förderung der Reflexionskompetenz besteht darin, dass die Reflexionen der meisten Studierenden bloße Beschreibungen der Praxis darstellen und nicht zu einer kritischen Bewertung oder Neuformulierung ihres Verständnisses führen. Leijen et al. (2014) haben deshalb in ihrer Studie ein angeleitetes Reflexionsverfahren entwickelt, das angehende Lehrerinnen und Lehrer dabei unterstützen soll, Wissen auf der Grundlage ihrer praktischen Erfahrungen zu entwickeln und mit theoretischem Wissen zu verknüpfen. Die Studierenden ($n = 22$) hatten unterschiedliche Vorerfahrungen und führten das Verfahren an einer Universität in Estland durch. In der ersten Phase wurde eine Unterrichtsstunde der Studierenden aufgezeichnet. Im Anschluss erhielten die Studierenden die Aufgabe, sich auf zwei Punkte zu konzentrieren. Ein Punkt sollte positiver Natur sein, der andere ein negativer Aspekt der Stunde, mit dem Ziel sich weiterentwickeln zu wollen. In Phase zwei reflektierten die Studierenden anhand von folgenden Leitfragen:

1. Was passiert an dieser Stelle? Was können Sie an sich selbst beobachten oder sogar hören? Was machen/sagen Ihre Schülerinnen und Schüler? Gibt es einen Zusammenhang zwischen dem, was Sie tun, und dem, was die Schülerinnen und Schüler tun?
2. Warum glauben Sie, dass das passiert ist? Welches Verhalten der Schülerinnen und Schüler wird durch Ihr Verhalten verursacht? Welches Verhalten von Ihnen wird durch das Verhalten der Schülerinnen und Schüler verursacht? Was macht den Vorfall für Sie zu einem kritischen Vorfall?
3. Wenn sie den Vorfall mit der Theorie in Beziehung setzen: Welche Lehrerrolle nehmen Sie ein? Wie unterstützt die Literatur Ihre Erklärung unter 2 (im Falle eines positiven Vorfalls)? Welche Vorschläge bietet die Literatur zur Lösung Ihres Problems (im Falle eines negativen Vorfalls)?
4. Was haben Sie von dieser Veranstaltung bisher gelernt? Wie werden Sie diese Dinge umsetzen?

Diese Phase wurde in drei unterschiedlichen Gruppen durchgeführt: mit einem Supervisor ($N = 7$), mit einem Peer-Studenten ($N = 6$) und allein ($N = 9$). Diese Einteilung konnten die Studierenden je nach Vorliebe selbst wählen und anschließend mündlich reflektieren.

In Phase drei fanden die schriftlichen Reflexionen ein bis zwei Wochen nach der mündlichen Reflexion von Phase zwei statt. Die Studierenden orientierten sich an den folgenden Leitfragen:

1. Wenn Sie den Vorfall mit der Theorie in Beziehung setzen: Welche Lehrerrolle nehmen Sie ein? Wie unterstützt die Literatur Ihre Erklärung unter 2 (im Falle eines positiven Vorfalls)? Welche Vorschläge bietet die Literatur zur Lösung Ihres Problems (im Falle eines negativen Vorfalls)?
2. Wie werden Sie zukünftig reagieren? Was werden Sie in Zukunft in Bezug auf diesen Vorfall unternehmen? Welches persönliche Konzept liegt Ihrer Handlungsentscheidung zugrunde?

3. Wie werden Sie die Dinge nutzen, die Sie in dieser Veranstaltung gelernt haben?

Diese empirische Studie konzentrierte sich auf die Erfahrungen der Lehrkräfte und das Feedback zum Verfahren der geführten Reflexion. Die Ergebnisse zeigten, dass der wahrgenommene Wert des Reflexionsverfahrens mit den früheren pädagogischen Erfahrungen der Lehrkräfte korreliert. Fast jeder bereits erfahrenere Studierende gab an, dass die einzelnen Phasen der Reflexion zu einer tieferen Analyse der ausgewählten Instanzen beigetragen haben. Ebenso erwähnten viele, dass mit jeder Phase der Reflexion neue Ideen entstanden sind. Darüber hinaus wiesen die meisten Studierenden darauf hin, dass die Fragen sie zu einem anderen Denken anregten, insbesondere die Analysen aus Schülersicht wurden als wichtig erachtet. Die Ergebnisse bezüglich der Reflexionen der Studierenden am Anfang des Studiums unterschieden sich deutlich. Die meisten Studierenden konnten nicht aufzeigen, wie ihnen das Reflexionsverfahren geholfen hat. Das wichtigste Feedback war das des Supervisors. Alle erklärten, dass dieser die Analyse von Emotionen und Gedanken unterstützt hat, die nicht aus dem Video oder aus den Reflexionen hervorgegangen waren. Nach Ansicht der Mehrheit der Studierenden am Anfang des Studiums sind die selbstbezogenen Gefühle und Gedanken die wichtigsten Aspekte ihrer Reflexionen. Mehr als die Hälfte der Studierenden, die allein die mündliche Reflexion durchführten, beschrieben es als schwierig, ohne Hilfsmittel zu reflektieren. Daher hatten sie das Gefühl, dass sie nicht so viel Nutzen aus dem Verfahren ziehen konnten wie diejenigen, die in Peers oder mit Supervisor reflektieren. So konnten abschließend fast alle erfahrenen Studierenden erklären, dass die angeleitete Reflexion ihnen bei der Überlegung half, inwieweit sie ihre persönlichen Ziele erreicht hatten. Die meisten der unerfahrenen Studierenden fanden das Verfahren der geführten Reflexion aufgrund fehlender Vorkenntnisse nicht vorteilhaft. Da dem Auftrag keine praktischen Erfahrungen folgten, hatten sie auch keine praktischen Möglichkeiten, ihre Unterrichtspraxis zu verbessern. Alle Studierenden gaben an, dass sie zumindest teilweise mit den Fragen nach der Theorie zu kämpfen hatten. Viele Studierende wiesen darauf hin, dass sie nicht wussten, wo sie die Literatur und die theoretischen Grundlagen finden würden. Andere gaben an, dass sie über ein allgemeines Wissen verfügen würden, es aber nicht auf bestimmte Theorien beziehen konnten. Darüber hinaus nannten viele fehlende pädagogische Kenntnisse, um über die ausgewählten Fälle nachzudenken.

Simone Abels (2011) untersuchte unter anderem das Reflexionsniveau und dessen Zusammenhang mit der Berufsidentität einerseits und einem demokratieförderlichen Naturwissenschaftsunterricht andererseits. Das Manual erreichte in der sechsten Überarbeitung Interraterreliabilitäten (prozentuale Übereinstimmung im Durchschnitt der vier Kodierertandems) zwischen 44 Prozent und 97 Prozent. Von den 39 Studierenden konnten vier ein hohes Reflexionsniveau erreichen, 12 das mittlere, 19 ein niedriges und vier konnten kein ausreichendes Niveau erreichen. Abels (2011) fand für das mittlere und hohe Niveau der Reflexionskompetenz nur tendenziell einen Zusammenhang zwischen Berufsidentität als Pädagoge und Reflexionskompetenz. In Bezug auf einen demokratieförderlichen Naturwissenschaftsunterricht zeigte sich: Je reflexiver der Studierende war, desto eher vertrat er die normativen Überzeugungen eines demokratieförderlichen Naturwissenschaftsunterrichts und stellte den durchgeführten Unterricht in diesem Sinne dar.

Basierend auf den Modellen von Korthagen und Kessels (1999) und Windt und Lenske (2015b) entwickelten Nowak, Ackermann und Borowski (2018) in Zusammenarbeit mit

Kempin, Kulgemeyer und Schecker (2019) ihr Modell zur Einschätzung von Selbst- und Fremdrelexionen von Fachunterricht (Abbildung 23).

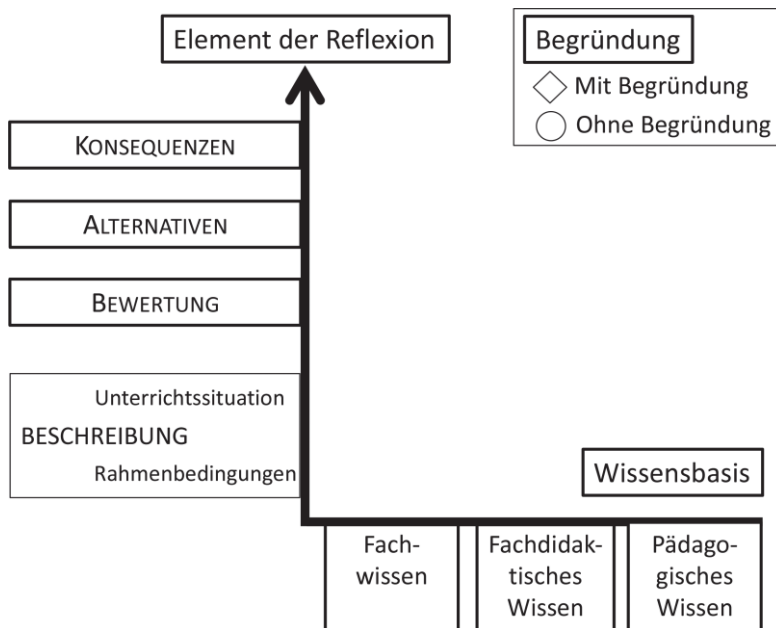


Abbildung 23: Modell zur Einschätzung von Selbst- und Fremdrelexionen von Fachunterricht (Nowak et al., 2018, S. 221)

Eine vollständige Reflexion beinhaltet hierbei fünf Elemente. Zunächst beginnt sie mit einer Beschreibung der Rahmenbedingungen. Diese soll die Ziele der Stunde, strukturelle Gegebenheiten und das Vorwissen der Schülerinnen und Schüler umfassen. Es folgt im nächsten Element eine möglichst objektive Beschreibung der Unterrichtssituation und im dritten Element eine Bewertung. Der Unterschied zwischen Alternativen und Konsequenzen wird im direkten Zusammenhang mit der Unterrichtssituation gezogen. Konsequenzen werden dabei als langfristige Ziele eingeschätzt. Jede Phase kann mit oder ohne Begründung vorliegen. Die Reflexionstiefe ist unterteilt in die drei Bereiche pädagogisches Wissen, fachdidaktisches Wissen und Fachwissen. Dabei werden die Reflexionen als strukturiert und unstrukturiert bewertet. Die Reflexionen gelten in diesem Zusammenhang als strukturiert, wenn ein Element geordnet auf das andere folgt, also alle fünf Phasen gefunden werden können. Um dies zu ermöglichen, wurden Leitfragen zur Unterstützung entwickelt (Abbildung 24).

LEITFRAGEN ZUR SELBSTREFLEXION VON UNTERRICHT

KONSEQUENZEN	→ Was ziehe ich aus der heutigen Stunde für meine persönliche Entwicklung als Lehrperson und warum?
ALTERNATIVEN	→ Was hätte ich anders machen können und warum wäre dieses Handeln erfolgreicher gewesen? → Oder warum würde ich es genauso wieder machen?
BEWERTUNG	→ Was war erwartungsgemäß? → Was war besser als erwartet? Und warum? → Was war schlechter als erwartet?
BESCHREIBUNG	→ Was habe ich gemacht? → Was haben die SuS gemacht? → Was ist passiert?
RAHMENBEDINGUNGEN	→ Warum habe ich die Methode/das Experiment/... ausgewählt? Was war mein Ziel? → Welche besonderen Aspekte habe ich bei der Planung bedacht?



Abbildung 24: Leitfragen zur Selbstreflexion (Nowak et al., 2018, S. 223)

Erste Einblicke in die Auswertung der Daten und die Ergebnisse gibt Kempin et al. (2019). Hier erhalten die Studierenden Unterrichtsausschnitte und zugehörige Prompts, welche die Reflexionspunkte vorgeben. Dies soll eine authentische Situation schaffen, in der die Reflexionsfähigkeit der Studierenden zum Ausdruck gebracht werden kann. Die Reflexionen werden wie folgt bewertet: Eine Beschreibung erhält einen Punkt, die Bewertung zwei, eine begründete Bewertung drei, eine Alternative vier Punkte, eine begründete Alternative fünf und eine Konsequenz sechs Punkte. Ist diese zudem begründet, erhält der Proband bzw. die Probandin sieben Punkte. Bei der Auswertung von $N = 69$ Studierenden konnte im Bereich der fachlichen und fachdidaktischen Wissensbasis eine nicht-signifikante Abnahme von Messzeitpunkt 1 auf Messzeitpunkt 2 verzeichnet werden. Im Bereich der erziehungswissenschaftlichen Wissensbasis wurde hingegen ein signifikanter Zuwachs ($d = 0.809$) erreicht. Als Grund wird die im Praxissemester evtl. vorliegende Verlagerung des Schwerpunktes auf erziehungswissenschaftliche Themen aufgeführt. In den Tests zu den Bereichen Fachwissen, fachdidaktisches Wissen und Erziehungswissenschaft konnte in allen drei ein signifikanter Wissenszuwachs mit mittleren Effekten (FW: $d = 0.472$, FDW: $d = 0.549$, EW: $d = 0.779$) verzeichnet werden.

Wehner und Weber (2018) erfassten in ihrer Pilotstudie die Reflexionskompetenz von Lehramtsstudierenden anhand von Fallvignetten. Hierbei werden reale oder fiktive Situationen dargestellt. In Anlehnung an Leonhard und Rihm (2011) wurden vier fiktive, aber plausible Fallvignetten entwickelt, die Bezug auf den Umgang mit Heterogenität in spezifischen Problemsituationen nehmen. Die Auswertung der Reflexionen erfolgte in Anlehnung an die Reflexionsebenen nach Bräuer und Schindler (2011) (Kapitel 3.4.1). Die Bearbeitung erforderte weder fachspezifische Kenntnisse noch diagnostisches Wissen. Um den Einfluss der Aufgabenstellung zu untersuchen, wurden drei unterschiedliche Instruktionen formuliert, die sich in den Operatoren und der Strukturierung unterschieden.

In Bezug auf die Aufgabenstellung konnten die bisherigen Ergebnisse der Forschung repliziert werden. So gehen strukturierte Instruktionen einher mit differenzierteren Ausführungen. Bei weniger strukturierten Aufgaben ist dies vom Operator unabhängig (Bräuer, 2009). Der Bezug zu wissenschaftlichen Theorien findet bei den Beurteilungen und Bewertungen kaum bzw. nur ansatzweise statt. Es konnten jedoch umfangreiche Ausführungen von Handlungsalternativen erzielt werden, was die Autoren auf das Anknüpfen an Erfahrungswissen zurückführen. Dies ist jedoch geprägt durch subjektive Theorien und deshalb wenig wissenschaftlich. Die unterschiedlichen Vignetten erreichten differenzierte Ergebnisse. Die Autoren führen dies auf den Inhalt zurück. So steht in einer Vignette die Lehrkraft im Mittelpunkt, mit welcher somit die Identifikation leichter fällt. Dies könnte zu den erfolgten Beurteilungen bzw. Bewertungen (Ebene 3) geführt haben. In der zweiten Vignette ist die Protagonistin eine störende Schülerin, was einen Perspektivwechsel erfordert. Dies ruft möglicherweise im ersten Schritt eher die eingetretene Interpretation (Ebene 2) hervor. Geplant für die Hauptstudie sind der ausschließliche Einsatz hochstrukturierter Fallvignetten und eine vertiefende Auswertung aller Kategorien, auch im Hinblick auf argumentative Zusammenhänge zwischen den einzelnen Ebenen. Die Entwicklung der Reflexionskompetenz wird anhand von Vergleichen zwischen Prä- und Posterhebungen untersucht werden.

Die Grundidee der Studie von Guardiera, Podlich und Reimer (2018) bestand darin, ein Lehrkonzept zu entwickeln, welches Studierenden in Vorbereitung auf das Praxissemester ermöglicht, vorgefasste Theorien und Haltungen in Bezug auf Unterricht mithilfe forschend angelegter Lernprozesse zu reflektieren und aufzubrechen. Ziel war es, einen Perspektivenwechsel zu ermöglichen und auch die Unterrichtswirklichkeit und wissenschaftliches Wissen in einer forschenden Grundhaltung aufeinander zu beziehen. An der Studie nahmen 37 Studierende des ersten Mastersemesters vollständig teil. Gruppenintern wurden im Regelfall 6er-Peergroups gebildet, um Feedback zu ermöglichen. Zunächst wurde mithilfe eines Fragebogens (18 Items) die Haltung zu (inklusive) Sportunterricht auf einer 5-stufigen Likert-Skala erfasst. Dieser wurde nach zehn Wochen wiederholt, um mögliche Veränderungen am Ende der Lehrveranstaltung abzubilden. Um Reflexionsprozesse gezielt anzustoßen, wurde fallorientiertes Arbeiten als geeignete Methode eingesetzt. In der wiederholten und intensiven Auseinandersetzung mit Fällen sollten zunächst Distanz zu Erstinterpretationen aufgebaut, das Deutungsrepertoire erweitert und Urteilsfähigkeit und Selbstreflektiertheit erweitert werden (Luesebrink & Grimminger, 2014). Als Vortest erhielten die Studierenden am Anfang der Lehrveranstaltung einen Projektfall, welcher mittels schriftlicher Analyse das vorhandene Reflexionsvermögen sowie das Reflexionsniveau erfassen sollte. Nach Ablauf der zehn Wochen wurde dieser Fall erneut verwendet. Zur Auswertung zog man ein deduktiv erstelltes Kategoriensystem heran, welches im Verlauf der Analyse induktiv ergänzt wurde. Kern der Untersuchung bildeten die Perspektivenübernahme, inwiefern wissenschaftliche Theoriebestände einbezogen wurden, und die Qualität erkennbarer Reflexionsprozesse. Letztere wurde auf Basis der von Wyss (2013) aufgestellten Qualitätsstufen der Reflexionsfähigkeit entwickelt (Kapitel 3.4.1). Zu bemerken bleibt, dass unter Perspektivenübernahme in dieser Studie die Reflexionsbreite verstanden wird. Als Perspektiven wurden u. a. folgende Kategorien bezeichnet: Classroom Management, Leistungsbeurteilung, Heterogenität, Lehrerverhalten, Klima, Unterrichtsstörung, etc. Ergebnisse konnten aufgrund der geringen Stichprobe nur rein deskriptiv berichtet werden, ergaben jedoch für einige Items des Fragebogens eine deutliche Verschiebung im Antwortverhalten und zeigten die Tendenz einer gesteigerten

Sensibilisierung für den Umgang mit und die Besonderheiten von heterogenen Lerngruppen. Bei der Reflexionsbreite (hier Perspektiven genannt) wurden sowohl im Vor- als auch im Nachtest sieben Kategorien gewählt. Insgesamt blieben nur sechs der 37 Probanden bei den ursprünglichen Breitenpunkten. 16 Teilnehmer und Teilnehmerinnen fügten zu ihren ursprünglichen weitere hinzu und 15 Probanden wählten gänzlich andere Punkte. Die Auswertung der Reflexionstiefe ist nicht eindeutig. Die Verwendung von Theoriewissen bei der Fallanalyse konnte von neun im Vortest auf alle Probanden, also 37, gesteigert werden. Jedoch sei angemerkt, dass nicht alle Teilnehmer dieses Wissen fallbezogen erläuterten, so dass eine bloße Benennung nicht unbedingt auf Wissen schließen lässt, sondern lediglich auf Nennung von Schlagwörtern reduziert wird. Eine Steigerung der Niveaustufen wurde nicht erreicht. So verzeichnete man sowohl im Vortest als auch im Nachtest die Reflexionsstufen 2 und 3 (persönliche Erfahrungsperspektive und Hinzunahme theoretischen Wissens). Die Reflexionsstufen 4 und 5 wurden sowohl im Vor- als auch im Nachtest nur von wenigen Einzelfällen erreicht.

Szogs, Krüger und Korneck (2019) untersuchten die Güte kollegialer Reflexionsgespräche und den Einfluss auf die direkte Verbesserung der Unterrichtsqualität im Fach Physik. Hierfür planten angehende Physiklehrkräfte im Microteaching-Seminar Unterrichtsminiaturen von zwölf Minuten, welche sie innerhalb eines Tages zweimal unterrichteten. Nach jeweils 5 Unterrichtsminiaturen erfolgte eine kollegiale strukturierte Reflexion. Diese beinhaltete zunächst eine Reflexion der jeweiligen Lehrkraft und im Anschluss eine kollegiale Peer-Reflexion inklusive zwei Experten mit dem Ziel, geeignete Handlungsalternativen zu entwickeln und so den Unterricht der zweiten Stunde zu verbessern. Am Ende wiederholte die Lehrkraft zentrale Ergebnisse der Reflexion und beschrieb geplante Änderungen der Stunde. Ausgewertet wurden sowohl die videografierten Unterrichtsstunden, als auch die Reflexionsgespräche. Kodiert wurde hierbei hinsichtlich des Inhalts (Unterrichtsqualitätsmerkmale, Professionalisierung), der Tiefe ((Rück-)Fragen, Beschreiben, Interpretieren, positive Rückmeldung, negative Rückmeldung, kritische Rückmeldung, Generierung von Handlungsalternativen) und weiterer Aspekte, wie Perspektivenwechsel und Begründungen. Es zeigte sich ein Anstieg der Unterrichtsqualität, wenn der Anteil an Aussagen mit interpretativem Charakter und der Anteil an Aussagen mit Begründungen erhöht war. Ein hoher Anteil von deskriptiven Aussagen führte zu einer Verschlechterung.

4. Forschungsziele, Forschungsfragen und Hypothesen

Das Kapitel 4 unterteilt sich in die Unterkapitel „Ziele der Arbeit“ und „Forschungsfragen und Hypothesen“. Nach Vorstellung der sechs Forschungsziele werden in Kapitel 4.2 die Forschungsfragen geklärt und die Hypothesen zugeordnet.

4.1. Ziele der Arbeit

Abgeleitet aus den Forschungsdesideraten zum Thema Reflexionskompetenz (Kapitel 3) und den widersprüchlichen Ergebnissen zur Möglichkeit, diese zu fördern (Kapitel 3.4.2), ergeben sich folgende Ziele dieser Arbeit:

- *Entwicklung eines universitären Seminars zur Förderung der Reflexionskompetenz von Chemielehramtsstudierenden*

Als Fördermaßnahmen werden Studierende ihre geplanten Unterrichtsstunden in einem Lehr-Lern-Labor (Kapitel 3.3.1) halten. Diese werden videografiert (Kapitel 3.3.2) und im Anschluss gemäß der Learning Study (Kapitel 3.3.3) wiederholt. Als weitere Unterstützungsmaßnahme wird Feedback (Kapitel 3.3.4) angeboten. Auf genauere Konzeptionen des Seminars wird in Kapitel 5 eingegangen.

- *Klärung des Zusammenhangs zwischen dem Fachwissen, dem fachdidaktischen Wissen und der Selbstwirksamkeitserwartung von Chemielehramtsstudierenden und der Entwicklung der Reflexionskompetenz*

Ergänzend wird der Zusammenhang zwischen der Entwicklung der Reflexionskompetenz und weiteren Aspekten der professionellen Handlungskompetenz von Lehramtsstudierenden untersucht (Kapitel 3.2.1).

- *Entwicklung eines Tests zur Erfassung des Wissens zum Thema Reflexionskompetenz*

In Anlehnung an die Definition des Begriffs Reflexion nach Wyss (2013) (Kapitel 3.1.1), sollen die schriftlichen Reflexionen mithilfe eines selbst entwickelten Kodiermanuals analysiert werden. Dieses berücksichtigt Aspekte zur Reflexionstiefe ebenso wie Aspekte zur Reflexionsbreite (Kapitel 3.4.1).

- *Entwicklung und Evaluation eines Kodiermanuals zur Erfassung der Reflexionstiefe und -breite*

Das Vorgehen bei der Entwicklung des Kodiermanuals ist in Kapitel 6.1.5 umfassend beschrieben. Das gesamte Manual ist auf CD im Anhang zu finden.

- *Vergleich des Zuwachses an Reflexionskompetenz bei Feedback vs. Selbstreflexion vs. Kontrollgruppe*

Um ein möglichst effektives Seminarkonzept zu entwickeln, werden die Studierenden unterschiedlich gefördert.

- *Aufklärung des Zusammenhangs zwischen einer Zunahme der Reflexionskompetenz von Chemielehramtsstudierenden und ihrem Handeln im Unterricht bzw. dem Lernzuwachs der Schülerinnen und Schüler*

Abschließend wird der Zusammenhang zwischen Reflexionskompetenz und Unterrichtswirksamkeit, welche über das Handeln im Unterricht sowie den Lernzuwachs der Schülerinnen und Schüler bestimmt wird, erforscht (Kapitel 2).

4.2. Forschungsfragen und Hypothesen

Im Zentrum dieser Studie standen zwei grundlegende Forschungsfragen:

F1: Wie kann ein universitäres Seminar gestaltet werden, um die Reflexionskompetenz von Chemielehramtsstudierenden zu fördern?

Ausgehend von den in Kapitel 3.1.1 erwähnten Definitionen wurden die verschiedenen Modelle zum Konstrukt Reflexion (Kapitel 3.1.2) thematisiert und ein zyklischer Ablauf des Seminars gewählt. Ebenso wurden verschiedene allgemeine Möglichkeiten der Förderung der Reflexionskompetenz umgesetzt, wie das Unterrichten in Lehr-Lern-Laboren oder die Aufzeichnungen der gehaltenen Stunden mittels Videografie mit anschließender Reflexion (Kapitel 3.3.2) und Feedback. So kann gefolgert werden:

H1 Durch das universitäre Seminar wird die Reflexionskompetenz der Chemielehramtsstudierenden gesteigert.

Als eine geeignete Methode zur Messung der Entwicklung der Reflexionskompetenz gelten schriftliche Berichte über Unterricht (Kapitel 3.4). So war es Ziel, ein für diese Studie passendes Manual zu entwickeln und zu evaluieren, welches die Reflexionskompetenz in den schriftlichen Überlegungen der Studierenden objektiv, reliabel und valide misst. Hierbei wurden die in verschiedenen Studien gemachten Erfahrungen eingearbeitet und umgesetzt (Kapitel 3.4.1).

H2 Die Reflexionskompetenz der Studierenden lässt sich mithilfe des entwickelten Kodiermanuals objektiv, reliabel und valide erfassen.

Eng mit der Förderung der Reflexionskompetenz verknüpft gilt die Professionalisierung in Bezug auf den Lehrberuf (Kapitel 3.2.1). Hierunter wird unter anderem der Lernzuwachs des fachdidaktischen Wissens, des Fachwissens und der Selbstwirksamkeit verstanden. Somit gilt es, folgende Frage zu klären:

F2: Welcher Zusammenhang besteht zwischen der Förderung der Reflexionskompetenz und dem fachdidaktischen Wissen, dem Fachwissen und der Selbstwirksamkeitserwartung der Chemielehramtsstudierenden und dem Lernzuwachs der Schülerinnen und Schüler?

Die Professionalisierung im Lehrberuf wird mit der Steigerung der Reflexionskompetenz in Zusammenhang gebracht (Kapitel 3.2.1). Es werden mögliche positive Auswirkungen von Reflexion auf das fachdidaktische Wissen angenommen, sowie des zyklischen Seminarverlaufs auf das Fachwissen. Somit gilt:

H3a Mittels des universitären Seminars lassen sich das experimentell-fachdidaktische Wissen und das Fachwissen zum Themenfeld Stofftrennung und -identifikation der Studierenden steigern.

H3b Der Lernzuwachs bezüglich der Reflexionskompetenz korreliert positiv mit dem experimentell-fachdidaktischen Wissen und dem Fachwissen zum Themenfeld Stofftrennung und -identifikation

Es ist anzunehmen, dass ein videobasierter Reflexions- und Feedbackzirkel einen motivierenden und positiven Effekt auf die allgemeine Selbstwirksamkeitserwartung der Studierenden haben kann. Ebenso können Zielorientierungen und fachspezifische Überzeugungen, vertieft und/oder gefiltert bzw. reflektiert und so das pPCK weiterentwickelt werden (Kapitel 3.2.1). So lässt sich folgende Hypothese aufstellen:

H4 Der Lernzuwachs bezüglich der Reflexionskompetenz der Studierenden korreliert positiv mit ...

H4.1 ... ihren Zielorientierungen im Studium.

H4.2 ... ihren fachspezifischen Überzeugungen.

H4.3 ... ihren Selbstwirksamkeitserwartungen.

Ebenso begründet in einer hohen Reflexionskompetenz liegt das Wahrnehmen und Bewusstwerden der Auswirkungen des Handelns im Unterricht (Kapitel 2). In dieser Studie wird eine Steigerung der Unterrichtswirksamkeit mit einem Lernzuwachs der Schülerinnen und Schüler begründet.

H5 Der Lernzuwachs im Fachwissen der jeweils an den Messzeitpunkten unterrichteten Schülerinnen und Schüler steigert sich von Messzeitpunkt I über Messzeitpunkt II bis zu Messzeitpunkt III.

Es sind zahlreiche positive Auswirkungen von externem Feedback auf verschiedene Bereiche wie Lernzuwachs und Unterrichtsqualität belegt (Kapitel 2 und 3.3.4). In der vorliegenden Studie wurden zwei Interventionsgruppen verglichen, die entweder Feedback erhielten und anschließend reflektierten oder lediglich selbst reflektierten. Somit ergab sich folgende Hypothese:

H6 Die Leistungen (der residuale Lernzuwachs) der Feedbackgruppe übertreffen (übertrifft) die (den) der Selbstreflexionsgruppe und der Kontrollgruppe in den Bereichen ...

H 6.1 Reflexionskompetenz

H 6.2 Lernzuwachs im experimentell-fachdidaktischen Wissen

H 6.3 Lernzuwachs im Fachwissen

H 6.4 Selbstwirksamkeit

H 6.5 Unterrichtswirksamkeit

5. Methodik und Durchführung

Kapitel 5 unterteilt sich in vier Unterkapitel. Zunächst werden das Untersuchungsdesign und die Stichprobengröße dieser Studie vorgestellt. Anschließend folgt ein Überblick über das entwickelte chemiedidaktische Hauptseminar. Es folgen detailliertere Beschreibungen, wie der Zeitpunkt des Einsatzes der Testinstrumente und die exakte Vorstellung dieser. Ebenso werden Veränderungen nach der Auswertung der Pilotstudie berichtet. Im letzten Teilkapitel findet sich die Datenauswertung. Beginnend mit den Testgütekriterien und den verwendeten statistischen Verfahren wird zusätzlich auf die Auswertung der schriftlichen Reflexionen mittels des erstellten Kodiermanuals eingegangen.

5.1. Untersuchungsdesign und Stichprobengröße

Untersuchungsdesign

Das in dieser Studie entwickelte chemiedidaktische Hauptseminar richtete sich an Studierende des Lehramts an Gymnasien und Realschulen. Auf Seiten der Studierenden wurde dabei zunächst mit einer Vorerhebung durch Paper-Pencil-Tests begonnen, welche Reflexionswissen, fachdidaktisches Wissen, Fachwissen zum Thema Stofftrennung und -identifikation, Zielorientierungen, fachspezifische Überzeugungen und Selbstwirksamkeitserwartungen beinhalten. Ermittelt wurden des Weiteren personenbezogene Daten, wie Alter, Geschlecht, Abiturnote und Informationen über den Stand bzw. die Einschätzung der universitären Lehramtsausbildung.

In Anlehnung an das Projekt von Anthofer (2017) erhielten die Teilnehmerinnen und Teilnehmer ihr Stundenthema mit verpflichtendem Schülerversuch, eine theoretische Einheit zur Planung von Unterricht sowie Reflexion und bereiteten im Anschluss eine Unterrichtsstunde vor. Diese wurde an einem ersten Schülertag gehalten und videografiert, im Anschluss analysiert und schriftlich reflektiert, überarbeitet und erneut gehalten. Es erfolgte eine Unterteilung der Studierenden in drei Gruppen nach dem Matched-Pairs-Prinzip auf Basis der Vortestergebnisse des CK-Tests. Die erste Gruppe erhielt zusätzlich zum Video der eigenen gehaltenen Unterrichtsstunde Feedback eines trainierten Mitarbeiters, direkt vor der schriftlichen Reflexion. Gruppe zwei analysierte die Stunde lediglich mithilfe der Videos und mittels Selbstreflexion. Eine Kontrollgruppe beobachtete die ihnen zugeteilte Stunde der Schülertage der ersten Gruppe, erhielt die entsprechenden Videos und reflektierte im Anschluss die eigens geplante Stunde, ohne den Unterricht selbst zu halten. Über drei Messzeitpunkte wurden hierdurch sowohl der Reflexionskreislauf von Korthagen (Korthagen & Kessels, 1999) als auch das Prinzip der Learning Studies (Wood & Sithamparam, 2014) umgesetzt. Die Schülertage zwei und drei verliefen analog zu Schülertag eins und wurden mit anderen Klassen und den jeweils verbesserten bzw. überarbeiteten Stundenentwürfen durchgeführt. In einer Nacherhebung am Ende des Seminars wurden die Posttests und nach Ablauf von ca. zwei Monaten Follow-Up-Tests der bereits aufgeführten Wissensbereiche abgehalten. Die schriftlichen Reflexionen der einzelnen Gruppen zu den drei Messzeitpunkten wurden mithilfe des erarbeiteten Kodiermanuals ausgewertet.

Um die Entwicklung der Unterrichtsstunden anhand des Lernzuwachses der Realschülerinnen und -schüler zu messen, wurde ca. zwei Wochen vor dem Schülertag das Fachwissen in Form eines Paper-Pencil-Tests erhoben. Das ausgewählte Stoffthema entsprach dem Chemieanfangsunterricht, entweder im naturwissenschaftlich-technischen Zweig in

der achten Jahrgangsstufe oder in den nicht-naturwissenschaftlich-technischen Zweigen in der neunten Jahrgangsstufe gemäß den Lehrplänen für die Realschulen in Bayern. Es ist ebenso verortet im Profilbereich des bayerischen G8-Lehrplans der Gymnasien in Bayern. Es wurde davon ausgegangen, dass die Entwicklung des Lernzuwachses abhängig vom Fachinteresse und den kognitiven Fähigkeiten der Schülerinnen und Schüler sei, welche zusätzlich einmal an diesem ersten Messzeitpunkt ermittelt wurden. Zudem wurde so der Vergleich der unterschiedlichen Klassen in Bezug auf die Kontrollvariablen sichergestellt. Für die Einhaltung des Pre-Post-Follow-Up-Designs wurde nach jeder Unterrichtsstunde am jeweiligen Schülertag und ca. zwei Monate nach dem Besuch an der Universität Regensburg erneut das Fachwissen überprüft.

Stichprobengröße

Die Stichprobenumfangsplanung verlief unter Verwendung der Software G*Power (Kapitel 10.3.1) (Faul, Erdfelder, Lang & Buchner, 2007; Faul, Erdfelder, Buchner & Lang, 2009). Unter Voraussetzung einer einfaktoriellen Varianzanalyse mit drei Gruppen und drei Messwiederholungen sowie Annahmen großer Effektstärken von $f = 0.6$ (Orientierung an dem Projekt von Anthofer (2017), Cohen, 1988) ergab dies eine Stichprobengröße von insgesamt 33 Teilnehmern. Dies konnte in der Hauptstudie mit 40 Studierenden umgesetzt werden (Tabelle 25).

Das chemiedidaktische Hauptseminar wurde insgesamt über drei Semester an der Universität Regensburg durchgeführt. Da im Wintersemester 2017/18 nur neun Studierende den Kurs besuchten, wurde dieser im Sommersemester 2018 an zwei Terminen (Kurs A und Kurs B) angeboten.

Tabelle 25: Stichprobengröße – Studierende der Interventionsgruppen (IG) und der Kontrollgruppe (KG) bzw. Schülerinnen und Schüler (SuS)

	Zeitraum	Anzahl Studierende			Anzahl SuS	
		IG 1 Feed-back	IG 2 Selbst-reflexion	KG Kontrollgruppe	IG 1 Feed-back	IG 2 Selbst-reflexion
Pilotstudie	Sommersemester 2017	6	6	6	1. Klasse: 22 2. Klasse: 15 3. Klasse: 12	1. Klasse: 20 2. Klasse: 23 3. Klasse: 15
Hauptstudie 1. Teil	Wintersemester 2017/18	4	5	-	1. Klasse: 25 2. Klasse: 27 3. Klasse: 27	1. Klasse: 22 2. Klasse: 11 3. Klasse: 20
Hauptstudie 2. Teil	Sommersemester 2018 (Kurs A)	5	5	7	1. Klasse: 22 2. Klasse: 22 3. Klasse: 22	1. Klasse: 16 2. Klasse: 26 3. Klasse: 16

	Sommersemester 2018 (Kurs B)	5	5	4	1. Klasse: 21 2. Klasse: 18 3. Klasse: 22	1. Klasse: 22 2. Klasse: 30 3. Klasse: 24
	Summe	20	21	17	255	245

Insgesamt haben 58 Studierende und 500 Schülerinnen und Schüler an dieser Studie teilgenommen. Hiervon waren 20 Studierende in der Interventionsgruppe 1 (Feedback), 21 Studierende in der Interventionsgruppe 2 (Selbstreflexion) und 17 Studierende in der Kontrollgruppe.

5.2. Durchführung der Studie

Ablauf des chemiedidaktischen Hauptseminars im Überblick

Das an der Universität Regensburg abgehaltene chemiedidaktische Hauptseminar wurde im Sommersemester 2017 (Pilotierung), im Wintersemester 2017/18 (Hauptstudie; Teil 1) und im Sommersemester 2018 (Hauptstudie; Teil 2) durchgeführt. Die einzelnen Pflichttermine umfassten jeweils 135 Minuten und wurden wöchentlich abgehalten. Die folgende Abbildung verdeutlicht den groben Ablauf (Abbildung 25):

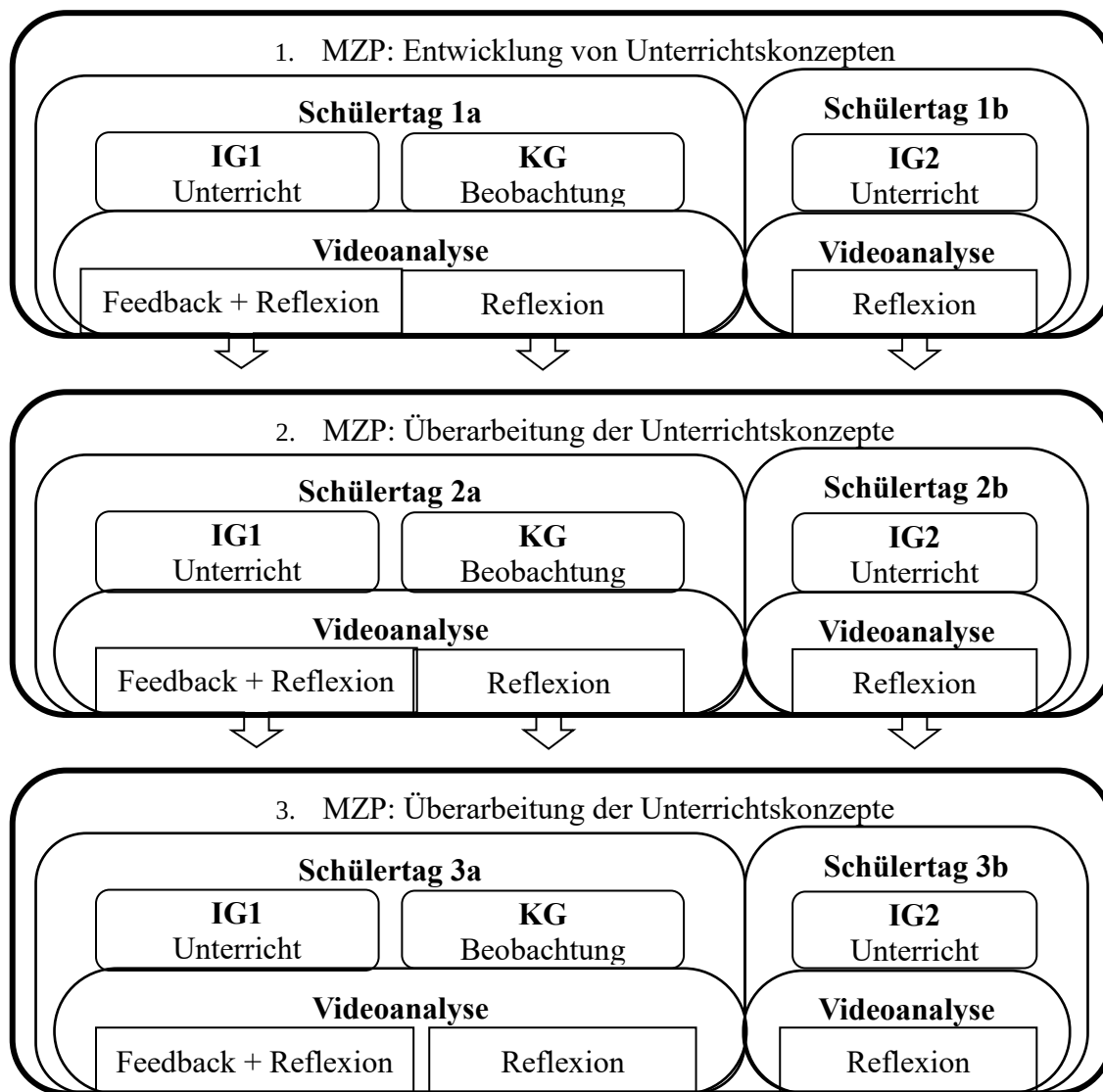


Abbildung 25: Ablauf und Durchführung des Hauptseminars im Überblick

Das Seminar folgte einem sich wiederholenden Kreislauf über drei Messzeitpunkte (MZP), basierend auf dem Prinzip der Learning Study (Kapitel 3.3.3) und dem Kreislauf der Reflexion im ALACT-Modell (Kapitel 3.1.2). Ausgehend von einem vorgegebenen Thema wurden Unterrichtsstunden konzipiert, an einem Schülertag erprobt und mithilfe der erstellten Unterrichtsvideos reflektiert, überarbeitet und abermals vor neuen Klassen gehalten. Dabei erfuhren die Studierenden unterschiedliche Treatments, welche im nächsten Absatz detailliert erläutert werden.

Ablauf des chemiedidaktischen Hauptseminars im Detail

1. Seminartermin: Testung

Der erste Seminartermin umfasste die Vorerhebungen des fachdidaktischen Wissens, des Fachwissens zum Thema Stofftrennung und -identifikation, der Zielorientierungen, der fachspezifischen Überzeugungen und der Selbstwirksamkeitserwartungen in Paper-Pencil-Form (Tabelle 26). Ebenso wurde das Wissen zum Thema Reflexion durch einen selbst erstellten Test abgefragt.

Tabelle 26: Durchgeführte Tests mit Quellenangaben

Test	Quelle
Fachdidaktisches Wissen	Backes, Sumfleth & Tepner, 2012
Fachwissen zum Thema Stofftrennung und -identifikation	Anthofer, 2017, S. 250–261
Zielorientierungen	Nitsche, Dickhäuser, Fasching & Dresel, 2011
Fachspezifische Überzeugungen	Lamprecht, 2011, S. 228
Selbstwirksamkeitserwartung	Meinhardt, Rabe & Krey, 2015; Rabe, Meinhardt & Krey, 2012

Da die Studierenden bis zu diesem Zeitpunkt wenig Erfahrung mit der Planung und Durchführung von Unterricht sammeln konnten, bezogen sich die Selbstwirksamkeitserwartungen auf die Vorbereitung und Durchführung von Experimenten und die didaktische Rekonstruktion unterrichtsrelevanter Inhalte.

2. Seminartermin: Theoretische Grundlagen und Unterrichtsplanung

Am zweiten Seminartermin wurde eine theoretische Grundlage zu den Themen Unterrichtsqualität und Reflexionskompetenz gelegt. Auf Basis der Ergebnisse des Fachwissenstests wurden die Studierenden mithilfe des Matched-Pairs-Systems in drei Gruppen unterteilt und jeder/jede Studierende entwickelte ein eigenes 40-minütiges Unterrichtskonzept zum Themenfeld Stofftrennung und -identifikation. Die Themenwahl begründet sich in der lehrplanunabhängigen Einsatzmöglichkeit, da es sich um Anfangsunterricht der achten Jahrgangsstufe des mathematische-naturwissenschaftlichen Zweiges bzw. der neunten Jahrgangsstufe der nicht-naturwissenschaftlich-technischen Zweige bayerischer Realschulen und ebenso der bayerischen Gymnasien handelt. Zusätzlich besteht die Möglichkeit zahlreicher geeigneter Schülerversuche, welche unterschiedliche Offenheitsgrade bieten, wodurch selbstgesteuertes Lernen möglich wird. Folgende Unterrichtsstunden und Versuche wurden vorgegeben (Tabelle 27):

Tabelle 27: Themen der Unterrichtseinheiten und zugehörige Experimente nach Anthofer (2017, S. 96)

Nr.	Thema	Experiment
1	Ableitung und Anwendung der Formel zur Bestimmung der Dichte	Bestimmung der Dichte von Magnesium- und Aluminium-Spitzen, Cent-Münzen und Metallstücken
2	Gewinnung von Kochsalz aus Steinsalz	Anwenden ausgewählter physikalischer Trennverfahren zur Reinigung von Steinsalz
3	Gewinnung der einzelnen Bestandteile eines Stoffgemisches	Anwenden mehrerer physikalischer Trennverfahren zur Gewinnung der einzelnen Bestandteile
4	Das Trennverfahren Destillation	Destillation von Rotwein oder Schnaps im Mikromaßstab und Nachweis des Destillats
5	Die Unterscheidung von Cola und Cola Light	Entfernen der braunen Cola-Färbung durch Aktivkohle und Durchführung der Fehling-Probe

6	Das Trennverfahren Chromatografie	Chromatografie am Beispiel von <i>Smarties</i> bzw. <i>M&Ms</i> unter Berücksichtigung der verschiedenen Farben

Die aufgeführte Reihenfolge entspricht der Pilotstudie. In der Hauptstudie 1 wurden aufgrund geringer Studierendenzahlen in der Feedbackgruppe die Stunden 3–6 gehalten. Die Selbstreflexionsgruppe hielt die Stunden 2–6. In Hauptstudie 2 wurden an allen Schülertagen die Stunden 2–6 gehalten. Die erste Stunde wurde aufgrund einer sehr physikalischen Umsetzung durch die Studierenden der Pilotstudie gestrichen. Alle Studierenden erhielten die gleichen Vorgaben. Dies waren das Thema der Stunde, ein Zeitrahmen von 40 Minuten und die grundsätzliche Struktur einer Unterrichtsstunde (Hinführung, Schülerexperiment und Auswertung). Die Studierenden konnten zur Unterrichtsplanung Schulbücher und Lehrerhandbücher nutzen, ebenso wie die Bibliothek der Chemiedidaktik der Universität Regensburg und das Internet.

3. Seminartermin: Unterrichtsplanung und Versuche

4. Seminartermin: Unterrichtsplanung und Versuche

In den darauffolgenden zwei Seminarterminen wurden die Unterrichtskonzepte weiterentwickelt, die Versuche getestet und wenn nötig überarbeitet und erneut geübt. Die weitere Abfolge war gruppenabhängig und wird durch Abbildung 25 genauer veranschaulicht.

5. Seminartermin: 1. Schülertag: IG 1 (Feedbackgruppe + Kontrollgruppe)

Beim nächsten Seminartermin wurde die erste Schülerklasse im Lehr-Lern-Labor der Universität Regensburg begrüßt (Abbildung 26). Nach einer kurzen allgemeinen Laboreinweisung und Sicherheitsbelehrung füllten die Schülerinnen und Schüler zunächst Paper-Pencil-Tests zum Fachinteresse und ihren kognitiven Fähigkeiten aus. Im Anschluss hielten die Studierenden der Feedbackgruppe ihre Stunden. Diese wurden von den Studierenden der Kontrollgruppe hospitiert, welche jeweils das gleiche Unterrichtsthema zugewiesen bekommen hatten.



Abbildung 26: Unterrichtsstunde im Lehr-Lern-Labor der Universität Regensburg

Nach jeder Unterrichtseinheit wurden die Schülerinnen und Schüler einem kurzen Wissenstest (bestehend aus 4–6 Items) zum Thema der Vorstunde unterzogen. Die Studierenden der Selbstreflexionsgruppe waren nicht anwesend.

6. Seminartermin: 1. Schülertag: IG 2 (Selbstreflexionsgruppe)

Der zweite Schülertag eine Woche später wurde entsprechend dem ersten Schülertag gestaltet. Hier fanden jedoch keine Hospitationen statt, es waren also weder die Studierenden der Feedbackgruppe noch die Studierenden der Kontrollgruppe anwesend.

7. Seminartermin: Reflexion und Überarbeitung

Der folgende Seminartermin war Pflichttermin für alle Teilnehmenden. Hier bekamen sie die Videoaufzeichnungen ihrer Unterrichtsstunde bzw. der hospitierten Unterrichtsstunde (Abbildung 27). Im Anschluss an diese 40 Minuten erhielten die Studierenden der IG 1 (Feedback) von einem trainierten Mitarbeiter Rückmeldung zu ihrer gehaltenen Stunde, reflektierten schließlich schriftlich und verbesserten ihre Unterrichtsstunden entsprechend. Sowohl IG 2 (Selbstreflexion) als auch die Kontrollgruppe begannen direkt über die Stunden zu reflektieren und verbesserten ebenso im Anschluss ihre Stundenentwürfe. In der Pilotstudie wurde die Zeit der Reflexionen auf 30 Minuten beschränkt. Allerdings empfanden dies die Studierenden als zu kurz. Da Schüpbach (2007) ausreichend Zeit als eine der Hauptbedingungen zum Gelingen der schriftlichen Reflexionen nennt, wurden die Reflexionen zeitlich nicht mehr beschränkt.



Abbildung 27: Analyse der Unterrichtsstunden mittels Videos im Lehr-Lern-Labor der Universität Regensburg

8. Seminartermin: 2. Schülertag: IG 1 (Feedbackgruppe + Kontrollgruppe)
9. Seminartermin: 2. Schülertag: IG 2 (Selbstreflexionsgruppe)
10. Seminartermin: Reflexion und Überarbeitung
11. Seminartermin: 3. Schülertag: IG 1 (Feedbackgruppe + Kontrollgruppe)
12. Seminartermin: 3. Schülertag: IG 2 (Selbstreflexionsgruppe)
13. Seminartermin: Reflexion

Die verbesserten Stunden wurden an den jeweiligen zweiten Schülertagen, entsprechend den ersten beiden Schülertagen, gehalten. Dieser Ablauf wurde zweimal wiederholt, so dass dieser insgesamt dreimal an drei Messzeitpunkten erfolgte.

14. Seminartermin: Testung

Im Anschluss an die dritten Schülertage und die letzten schriftlichen Reflexionen erfolgten zum nächsten Seminartermin die Posttests zum fachdidaktischen Wissen, zum Fachwissen, zu den Zielorientierungen, fachspezifischen Überzeugungen, den Selbstwirksamkeitserwartungen und zur Reflexionskompetenz.

Videografie

Die gehaltenen Unterrichtsstunden wurden sowohl in der Pilot- als auch in der Hauptstudie videografiert. Hierbei waren drei Kameras im Einsatz. Aus Lehrerperspektive wurde

die Totale von einer Kamera erfasst. Aus Schülerperspektive wurde ebenfalls eine Kamera zur Aufnahme des Gesamtspektrums hinten verwendet. Eine weitere Kamera fand Einsatz als Aktionskamera und folgte dem bzw. der unterrichtenden Studierenden. Ebenso wurden Mikrofone im Klassenzimmer installiert. Dabei zeichneten vier Mikrofone die Gespräche der Schülerinnen und Schüler auf. Ein weiteres Mikrofon wurde von den Studierenden am Körper getragen. Abbildung 28 verdeutlicht den Aufbau im Klassenzimmer.

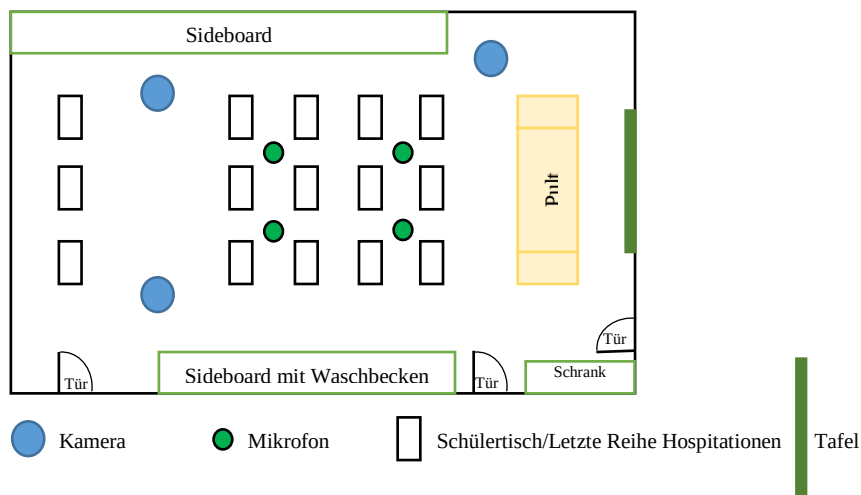


Abbildung 28: Position der Kameras und Mikrofone im Lehr-Lern-Labor

Vorgehen beim Feedback

Zunächst wurden die Unterrichtsstunden an den Schülertagen der IG 1 von dem Mitarbeiter hospitiert. Die zusätzlichen Videoaufzeichnungen aus drei unterschiedlichen Perspektiven machten Beobachtungen aus verschiedenen Blickwinkeln ebenso möglich wie die wiederholte Betrachtung von Schlüsselsituationen oder eventuell verpasster Aktionen des Stundenverlaufs. Unterstützend wurden die Tonaufnahmen zur Verfügung gestellt. So konnten durch die Schülermikrofone auch die Konversationen während der Gruppenarbeit nachvollzogen werden. Dadurch war eine umfassende Beobachtung sichergestellt. Dem Mitarbeiter waren hierbei die Beobachtungsaspekte ebenso wie die Beurteilungstendenzen stets klar (Beobachtungs- und Beurteilungsfehler sind nachzulesen unter anderem in Helmke, 2009). So war das zu beobachtende Verhalten eindeutig vordefiniert und ließ sich von anderem Verhalten abgrenzen. Voraussetzung war dabei die Kenntnis der dahinterliegenden Theorie, welche aufgrund einer gewissenhaften Einarbeitung des Mitarbeiters im Rahmen seiner Zulassungsarbeit zum Thema Feedback als gegeben vorausgesetzt werden konnte.

Die 30-minütigen Feedbacktermine begannen mit einem kurzen Überblick über den Ablauf des folgenden Gesprächs. Im ersten Feedbackgespräch wurden vorwiegend verhaltensnahe Aspekte sowie positive und verbesserungsfähige Punkte angesprochen. Im Anschluss wurden die Studierenden nach ihrer persönlichen Einschätzung der Stunde gefragt. Aus diesen Punkten wurden Verbesserungs- bzw. Änderungsmöglichkeiten herausgearbeitet, die am zweiten Schülertag umgesetzt werden konnten. Bei diesem praktischen Teil der Intervention sollten die Studierenden ihre Erfahrungen mit dem Feedback individuell verarbeiten. Der theoretische Abschnitt des Gesprächs enthielt ausgewählte Aspekte zum Themenfeld Unterrichtsfeedback. Dies betrifft unter anderem Punkte zum Ge-

ben und Nehmen von Feedback oder die Merkmale guten Unterrichts. Im zweiten Feedbackgespräch wurden lerntheoretische Grundlagen besprochen und Aspekte des selbst-gesteuerten Lernens. Im praktischen Teil wurde vermehrt auf die allgemeine Gestaltung der gehaltenen Unterrichtsstunde eingegangen, außerdem erfolgte eine Ergänzung um verhaltensnahe Aspekte. Abschließend wurden die Studierenden wiederum nach ihrer Sichtweise zum Unterrichtsverlauf befragt und gemeinsam Verbesserungsmöglichkeiten erarbeitet. Das dritte und letzte Feedbackgespräch verlief analog zum ersten mit verstärkten Bezügen zu den ersten beiden Unterrichtsstunden und den hierzu gegebenen Feedbackpunkten. In allen Gesprächen wurden zahlreiche Aspekte thematisiert, einschließlich etwaiger Punkte wie Überleitungen und Hinführungen zu Themen, Zeitmanagement, Körperhaltung, Gestik und Mimik, Sprachlautstärke, Bewegung im Raum, Präsentation von Lehrer-Experimenten, Umgang mit Missgeschicken, Fachwissen (Ergebnisse der Zulassungsarbeit und der genaue Ablauf nachzulesen bei Heber, 2018).

5.3. Erhebungsinstrumente

Kapitel 5.3 beinhaltet alle Testinstrumente und Fragebogen. Dies sind auf Seite der Studierenden der Test zum fachdidaktischen Wissen, der Fachwissenstest, der Studierendenfragebogen und der Reflexionskompetenztest. Ebenso wird das zur Auswertung der schriftlichen Reflexionen eingesetzte Kodiermanual vorgestellt. Seitens der Schülerinnen und Schüler wird der Wissenstest, der kognitive Fähigkeitstest und der Fragebogen zum Fachinteresse erläutert. Um einen reibungslosen und vergleichbaren Ablauf zu garantieren und um sicherzustellen, dass alle Probanden die Tests ohne Verwendung von Hilfsmitteln ausfüllen, war bei der Durchführung mindestens eine Aufsichtsperson anwesend. Die Testleiter können auf zwei Personen, eine davon die Leiterin dieser Studie, eingeschränkt werden. Auf die Testgütekriterien Objektivität, Reliabilität und Validität wird in Kapitel 5.4.1 genauer eingegangen.

5.3.1. Testinstrumente und Fragebogen – Studierende

Test zum fachdidaktischen Wissen

Das fachdidaktische Wissen wurde mithilfe eines von Backes et al. (2012) entwickelten Tests im Pre-Post-Follow-Up-Design abgefragt, bestehend aus insgesamt 19 Items, die sich jeweils zusammensetzten aus einer Situationsdarstellung, einer Bewertungsaufforderung und vier möglichen Antwortalternativen. Letztere sollen auf ihre Relevanz für die jeweilige Aufgabe auf einer sechsstufigen Likert-Skala bewertet werden. Die Situationen beschreiben konkrete Punkte während der Planung, Durchführung und Evaluation des Chemieunterrichts in der Schule mit Schwerpunkt auf Auswahl und Umsetzung von Experimenten im Unterricht, Schülervorstellungen und Modelleinsatz. Um sich besser in die verschiedenen Konstellationen hineinversetzen zu können, wurden die Situationen aus der Lehrerperspektive formuliert, welche fiktive Referendare hinsichtlich ihres Verhaltens einschätzen sollten. Diese Distanzierung fördert die Teilnahmebereitschaft und kann positive Auswirkungen auf das Professionswissen haben (Schwindt, 2008; Witner & Tepner, 2011). Zur Verdeutlichung wird Item 18 exemplarisch aufgeführt (Backes et al., 2012) (Abbildung 29):

Ihr Referendar möchte im Chemieanfangsunterricht selbstständig experimentieren lassen. Auf welche typischen Schwierigkeiten der SuS sollte er besonders achten?

Bitte bewerten Sie die Relevanz folgender Vorschläge. Geben Sie dazu bitte Noten von 1 („sehr gut/zutreffend“) bis 6 („ungenügend/unzutreffend“):

	1	2	3	4	5	6
SuS lenken sich beim Experimentieren häufig gegenseitig ab.	☐	☐	☐	☐	☐	☐
SuS machen in der Durchführung häufig manuelle Fehler.	☐	☐	☐	☐	☐	☐
SuS haben häufig Angst vor Chemikalien.	☐	☐	☐	☐	☐	☐
SuS halten sich häufig sehr lange mit der Dokumentation ihrer Arbeitsschritte auf.	☐	☐	☐	☐	☐	☐

Abbildung 29: Beispielitem aus dem Test zum fachdidaktischen Wissen – Studierende

Ausnahmen zu oben beschriebenem Aufbau stellen Item 6, in welchem die Studierenden aufgefordert werden, acht Schritte des experimentellen Erkenntnisgewinnungsprozesses in eine dem wissenschaftlichen Vorgehen entsprechende Reihenfolge zu bringen, und Item 19 dar, in welchem die Studierenden acht Aussagen den entsprechenden Schritten des experimentellen Erkenntnisgewinnungsprozesses zuordnen sollen. Exemplarisch wird Item 6 aufgeführt (Schmitt, 2016, S. 246) (Abbildung 30):

Bitte bringen Sie die folgenden Schritte des experimentellen Erkenntnisgewinnungsprozesses in eine dem wissenschaftlichen Vorgehen entsprechende Reihenfolge:

- a) Eine Hypothese aufstellen
- b) Eine Schlussfolgerung formulieren
- c) Ein Experiment beobachten
- d) Ein Experiment durchführen
- e) Versuchsergebnisse analysieren & interpretieren
- f) Deskriptive Ergebnisse darstellen
- g) Eine experimentelle Untersuchung planen
- h) Eine konkrete experimentell-prüfbare Fragestellung herausarbeiten

h → a → g → d → c → f → e → b

Abbildung 30: Beispielitem aus dem Test zum fachdidaktischen Wissen – Studierende

Die Bearbeitungsdauer umfasste je Messzeitpunkt 20 Minuten. Zur Bewertung der einzelnen Antwortalternativen wurde ein Expertenrating durchgeführt. Mithilfe paarweiser Vergleiche konnten so Reihenfolgen festgelegt werden. Die Einschätzungen der Studierenden wurden mit diesen verglichen und pro Übereinstimmung ein Punkt vergeben. Item 6 wurde bei richtiger Zuordnung aller Schritte mit zwei Punkten bewertet, ansonsten mit null Punkten. Bei Item 19 konnten pro richtiger Zuordnung 0.25 Punkte erzielt werden (Witner & Tepner, 2011).

Fachwissenstest

Der Fachwissenstest zum Thema Stofftrennung und -identifikation wurde von Anthofer (2017) entwickelt und setzte sich ursprünglich aus 26 Items mit jeweils fünf Antwortmöglichkeiten zusammen. Die ersten zehn Fragen waren auf dem Niveau der Sekundarstufe I, die Fragen elf bis achtzehn auf dem Niveau der Sekundarstufe II und die letzten acht Fragen wurden dem Wissenstand des Grundstudiums Chemie zugeordnet. Zusätzlich wurde darauf geachtet, deklaratives, prozedurales und konditionales Wissen abzufragen (Paris et al., 1983). Nach ausführlicher Validierung (Kapitel 5.4.1) im Projekt von Anthofer (2017) wurden für diese Studie 21 Items übernommen, darunter acht Items zur Sekundarstufe I, acht Items zur Sekundarstufe II und fünf Items auf Basis des Grundstudiums Chemie. Beispielhaft wird ein Item dieses Tests aufgeführt (Anthofer, 2017, S. 250) (Abbildung 31):

Bei welcher der folgenden Flüssigkeiten handelt es sich um eine Suspension?	
Erdöl	☐
Wandfarbe	☒
Mayonnaise	☐
Salatdressing aus Essig und Öl	☐
Milch	☐

Abbildung 31: Beispielitem aus dem Test zum Fachwissen – Studierende

Die Bearbeitungszeit betrug ca. 25 Minuten. Pro Frage existierte exakt eine richtige Antwortmöglichkeit, welche mit einem Punkt bewertet wurde. Bei falsch angekreuzten Aufgaben erhielten die Studierenden keine Punkte.

Studierendenfragebogen

Der eingesetzte Studierendenfragebogen wurde durch Anthofer (2017, S. 106–108) aus unterschiedlichen Quellen adaptiert (Tabelle 28), modifiziert und in dieser Studie übernommen. Er bestand aus vier Teilen, angefangen mit persönlichen Angaben wie Alter, Geschlecht, Abiturschnitt und Fragen zur vergangenen und aktuellen Ausbildung. Es folgten 16 Items zu Zielorientierung, 39 Items zu fachspezifischen Überzeugungen und 22 Items zur Selbstwirksamkeitserwartung. Die einzelnen Items wurden auf einer fünfstufigen Likert-Skala von den Studierenden im Pre-Post-Follow-Up-Design beantwortet.

Tabelle 28: Originalquellen des Studierendenfragebogens

Teil des Fragebogens	Originalquelle
Zielorientierungen ($N_{Items} = 16$)	Nitsche et al., 2011
Fachspezifische Überzeugungen „Chemie als Wissenschaft“ ($N_{Items} = 10$)	Lamprecht, 2011, S. 228

Fachspezifische Überzeugungen „Das Unterrichtsfach Chemie“ ($N_{Items} = 21$)	Lamprecht, 2011, S. 229–231
Fachspezifische Überzeugungen „Sozialformen im Chemieunterricht“ „Selbststeuerung beim Lösen von Problemstellungen und Durchführen von Experimenten“ ($N_{Items} = 8$)	Lamprecht, 2011, S. 233–234
Selbstwirksamkeitserwartungen „Planung von Experimenten“ ($N_{Items} = 5$) „Durchführung von Experimenten“ ($N_{Items} = 4$) „Didaktische Rekonstruktion“ ($N_{Items} = 8$) „Umsetzung der didaktischen Rekonstruktion“ ($N_{Items} = 5$)	Meinhardt et al., 2015; Rabe et al., 2012

Exemplarisch werden ausgewählte Items aus den verschiedenen Bereichen aufgeführt (Tabelle 29, Tabelle 30, Tabelle 31):

Tabelle 29: Beispielitem Zielorientierung

In meinem Studium strebe ich danach ...	sehr zu- treffend	zu- treffend	mit- tel	unzu- treffend	sehr unzu- treffend
... neue Ideen zur Vermittlung von Wissen in meinem Fach zu bekommen.	☐	☐	☐	☐	☐
... komplizierte Vorlesungen besser zu verstehen.	☐	☐	☐	☐	☐
... meine pädagogischen Kompetenzen zu verbessern.	☐	☐	☐	☐	☐

Tabelle 30: Beispielitem Fachspezifische Überzeugung

Lehren von Chemie in der Schule	sehr zu- treffend	zu- treffend	mit- tel	unzu- treffend	sehr unzu- treffend
Zu einem guten Chemieunterricht gehört in der Regel das experimentelle Arbeiten der Schülerinnen und Schüler (SuS).	☐	☐	☐	☐	☐
Es ist nicht ausreichend, Experimente lediglich in der Theorie zu behandeln.	☐	☐	☐	☐	☐
Eine Chemielehrkraft ist vorrangig Chemiker.	☐	☐	☐	☐	☐

Tabelle 31: Beispielitem Selbstwirksamkeitserwartungen

	sehr zu- treff- end	zu- treff- end	mit- tel	unzu- treff- end	sehr unzu- treffend
Es bereitet mir keine Probleme, zu einem Thema passende Experimente vorzubereiten.	☐	☐	☐	☐	☐
Ich kann ein Demonstrationsexperiment für Schülerinnen und Schüler nachvollziehbar durchführen, auch wenn es sich um eine sehr komplexe Versuchsanordnung handelt.	☐	☐	☐	☐	☐
Ich schaffe es, auch komplexe Themen der modernen Chemie für eine Unterrichtsstunde so zu vereinfachen, dass meine Schülerinnen und Schüler sie verstehen.	☐	☐	☐	☐	☐

Die Studierenden hatten für diese Befragung an allen drei Messzeitpunkten (Pre-Post-Follow-Up) 30 Minuten Zeit. Für die Berechnung der Punkte wurden zunächst die 25 Items recodiert. Im Anschluss wurden ein bis fünf Punkte aufsteigend der Likert-Skala zugeordnet und der Durchschnitt berechnet. Dann folgte ein Vergleich mit den Durchschnitten über die drei Messzeitpunkte.

Reflexionskompetenztest

Im Frühjahr 2017 wurde ein Reflexionskompetenztest bestehend aus 21 Items entwickelt. Diese Multiple-Choice Fragen umfassten jeweils fünf Antwortmöglichkeiten, von denen je eine richtige ist. Die Konstruktion begann mit der Festlegung der Ziele und der Zielgruppe. So wurde entschieden, einen Wissenstest mit sowohl deklarativen als auch prozeduralen Items zu gestalten. Die Zielgruppe war durch das chemiedidaktische Seminar klar und umfasste Studierende des Lehramts der Chemie für Gymnasien und Realschulen aus den Semestern 5–12. Ferner wurden die Fragen auf Basis von Anforderungsanalyse und Literatur erstellt (Thielsch, Lenzner & Melles, 2012). Bei der Konzeption der Fragen dienten folgende Grundregeln als Orientierungshilfe (Thielsch et al., 2012, S. 225–231):

- Formulierung leicht verständlicher Fragen
 - grammatikalisch richtig
 - Alltags- und nicht Wissenschaftssprache
 - sachlich und neutral (Converse & Presser, 1986, 10ff.)
 - Prüfung auf leichter verständliche Synonyme (Bsp.: Diskrepanz und Unterschied)
- Formulierung kurzer, aber nicht zu kurzer Items
- Vermeidung überflüssiger Fragen
 - Fragen, welche nur evtl. ausgewertet werden, sollten sehr sparsam verwendet werden (Bortz & Döring, 2007)
- Vermeidung von Suggestivfragen

- der Stil der Frage sollte die Befragten ermuntern, das zu antworten, was sie als richtig erachten (Bortz & Döring, 2007)
- Sinnvolle Strukturierung des gesamten Bogens
 - Fragen zum gleichen Thema innerhalb eines Frageblocks integrieren
 - vom Allgemeinen zum Speziellen sortieren (Bradburn, Sudman & Wansink, 2004)
 - nach dem sogenannten Eisbrecheritem folgt eine Itemreihung mit ansteigender Schwierigkeit
- nur eine Aussage je Frage
- Doppelfragen (Oder-Fragen) vermeiden (Fisseni, 2004)
- Vermeidung doppelter Verneinungen (Bühner, 2011)
- Einschränkung auf Begriffe mit klarer Bedeutung
- Vermeidung telegrafischer Kürzel
- syntaktische Struktur möglichst einfach halten
 - Passiv- in Aktivsätze überführen
 - Schachtelsätze vermeiden (Lenzner, Kaczmirek & Lenzner, 2010)
- Hervorhebung wichtiger Inhalte
 - Aber: zu häufige Verwendung ist kontraproduktiv (Bühner, 2011)
- Formulierung ichbezogen gestalten
- keine zu allgemeinen Formulierungen (Bortz & Döring, 2007; Bühner, 2011)
- Vermeidung von Items, welchen grundsätzlich zugestimmt/abgelehnt werden
- positiv und negativ formulierte Items für spätere Analysen merken
- für jedes Item prüfen, ob das Konstrukt abgebildet wird (Bühner, 2011)

Die Aufgabenzusammensetzung der Pilotstudie bestand aus elf deklarativen Fragen und zehn konditionalen Fragen. Nach der ersten Testung wurden Items mit einer geringeren Trennschärfe als 0.15 verworfen; so blieben neun deklarative Fragen und acht konditionale übrig. Letztere folgen einem sich wiederholenden Muster aus Arbeitsanweisung, Falldarstellung und schließlich Lösungsmöglichkeiten. Hierbei sollen die Probanden verschiedene fiktive Situationen den vier Reflexionsstufen von Hatton und Smith (1995) zuordnen. Exemplarisch werden auch hier zwei Beispielitems aus den verschiedenen Bereichen aufgeführt (Abbildung 32, Abbildung 33):

Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu: „Bei der Hypothesenbildung tat sich die Gruppe 3 sehr schwer. Hier wäre ein Tipp von mir wohl hilfreich gewesen.“	
„descriptive reflection“ (beschreibende Reflexion)	☒
„descriptive writing“ (Beschreibung)	☐
„critical reflection“ (kritische Reflexion)	☐
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☐

Abbildung 32: Beispielitem des Reflexionskompetenztests – konditionales Wissen

Das ALACT-Modell nach Korthagen beschreibt einen Reflexionskreislauf in folgenden fünf Phasen:	
1. Handlung	☒
2. Rückblick auf die Handlung	
3. Bewusstsein der essenziellen Aspekte	

4. Kreieren alternativer Handlungsformen 5. Veränderte Handlung	
1. Handlung 2. Rückblick auf die Handlung 3. Kreieren alternativer Handlungsformen 4. Bewusstsein der essenziellen Aspekte 5. Veränderte Handlung	☐
1. Handlung 2. Bewusstsein der essenziellen Aspekte 3. Rückblick auf die Handlung 4. Kreieren alternativer Handlungsformen 5. Veränderte Handlung	☐
1. Handlung 2. Kreieren alternativer Handlungsformen 3. Rückblick auf die Handlung 4. Bewusstsein der essenziellen Aspekte 5. Veränderte Handlung	☐
1. Handlung 2. Kreieren alternativer Handlungsformen 3. Bewusstsein der essenziellen Aspekte 4. Rückblick auf die Handlung 5. Veränderte Handlung	☐

Abbildung 33: Beispielitem des Reflexionskompetenztests – deklaratives Wissen

Die Bearbeitungsdauer betrug an allen drei Messzeitpunkten 25 Minuten. Pro Frage existierte eine richtige Antwortmöglichkeit, welche mit einem Punkt bewertet wurde. Bei falsch angekreuzten Aufgaben erhielten die Studierenden keine Punkte. Die Ergebnisse zu den Testgütekriterien und zu den Bearbeitungen der Probanden werden in den Kapiteln 6.1 und 7 ausführlich dargestellt. Der vollständige Test findet sich in Kapitel 10.2.

Kodiermanual

Im Rahmen einer Zulassungsarbeit wurde ein Kodiermanual zur Auswertung der schriftlichen Reflexionen mithilfe der Qualitativen Inhaltsanalyse entwickelt. Dies ist „eine empirische Methode zur systematischen, intersubjektiv nachvollziehbaren Beschreibung inhaltlicher und formaler Merkmale von Mitteilungen, meist mit dem Ziel einer darauf gestützten interpretativen Inferenz auf mitteilungsexterne Sachverhalte“ (Früh, 2015, S. 29). Mayring (2010) fasst die Aufgaben der Inhaltsanalyse wie folgt zusammen:

- „Kommunikation analysieren.
 - fixierte Kommunikation analysieren.
 - dabei systematisch vorgehen.
 - dabei also regelgeleitet vorgehen.
 - dabei auch theoriegeleitet vorgehen.
 - das Ziel verfolgen, Rückschlüsse auf bestimmte Aspekte der Kommunikation zu ziehen.“
- (Mayring, 2010, S. 13)

Das für die Zulassung entwickelte Manual konnte jedoch nicht ausreichend überzeugen. Die Werte lagen hierbei für die Reflexionsbreite bei $\kappa = .82$ (sehr gut), für die Reflexionstiefe bei lediglich $\kappa = .56$ (mäßig) und insgesamt bei $\kappa = .67$ (gut). So wurden vor allem

die Kategorien der Reflexionstiefe substanziell überarbeitet. Auch das Bewertungsschema der Reflexionsniveaus konnte für diese Studie nicht übernommen werden. Das zunächst hauptsächlich an Hatton und Smith (1995) sowie Abels (2011) angelehnte Manual teilte die Testpersonen durch die Verwendung verschiedener Formulierungen in unterschiedliche Niveaus der Reflexionskompetenz ein. So wurde zum Beispiel eine höhere Niveaustufe durch Perspektivenübernahme erreicht oder aber durch Literaturbezüge. Dies konnte aber in der Zulassungsarbeit nicht ausreichend überzeugen, da beispielsweise eine Reflexion mit geringem Inhalt lediglich aufgrund nur einer auftretenden Perspektivübernahme mit einem höheren Niveau bewertet wurde. Folglich ließ sich die Einordnung der Reflexionskompetenz der Probanden nicht ausreichend genau darstellen. Das Ziel bestand darin, die Bewertung unabhängig von Formulierungsstärken, -schwächen bzw. -vorlieben der Probanden durchzuführen. Konsens war schließlich, pro Reflexionsart (Bewertung, Rechtfertigung, Konsequenzen ziehen, Perspektivenübernahmen etc.) Punkte zu verteilen und die erreichten Summen in Niveaustufen einzuordnen. So konnte eine Testperson, welche zum Beispiel wenig Perspektivenübernahme zeigte, diese durch andere Stärken, wie etwa Literaturbezüge, ausgleichen. Als wichtig hierbei erschien ebenso die einheitliche Bepunktung aller Reflexionsarten, da vor allem die Probanden in der Kontrollgruppe vermehrt Verbesserungsvorschläge machten und weniger negative Bewertungen vornahmen. Genauere Ergebnisse dazu finden sich in Kapitel 7. So wurde eine unterschiedliche Bepunktung und somit Gewichtung der Reflexionsarten für diese Studie verworfen (Kempin et al., 2019; Nowak et al., 2018; Windt & Lenske, 2015b). Auch ein Durchlaufen eines immer gleichen Schemas von Reflexionsarten zum Erreichen eines hohen Niveaus konnte für diese Studie nicht überzeugen (Kempin et al., 2019). Die Probanden sollten ihre Reflexionen frei formulieren und nicht ein festes Schema, bestehend aus einem festen Ablauf, durchlaufen müssen, um etwaige Entwicklungen messen zu können. Folglich wurde grundsätzlich ein Punkt pro auftretender Reflexionsart verteilt und Formulierungsvorlieben somit nivelliert. Bei der Entwicklung des überarbeiteten Kategoriensystems fanden die Techniken qualitativer Inhaltsanalyse Beachtung (Mayring, 2010). Es wurde darauf geachtet, dass das Material in den Kommunikationszusammenhang eingebettet war. Aus diesem Grund wurde der Text stets innerhalb des Kontextes interpretiert und das Material auf seine Entstehung und Wirkung hin untersucht. Ebenso wurde, wie von Mayring gefordert, systematisch, das heißt theoriegeleitet und regelgeleitet, vorgegangen. Alle beschriebenen Techniken der qualitativen Inhaltsanalyse wurden Schritt für Schritt eingehalten. Die drei Grundformen des Interpretierens sind: Zusammenfassung, Explikation und Strukturierung. Im ersten Schritt wurden die schriftlichen Reflexionen der Pilotstudie systematisch auf das Wesentliche reduziert. Dies entspricht einer induktiven Kategorienbildung. Die Analyseeinheit wird bestimmt, die inhaltstragenden Textstellen werden paraphrasiert, die Paraphrasen unter einem festgelegten Abstraktionsniveau generalisiert, bedeutungsgleiche Paraphrasen gestrichen; darauf folgt die Bündelung von inhaltlich zusammenpassenden Paraphrasen, die Zusammenfassung zu einem Kategoriensystem sowie die Rücküberprüfung am Ausgangsmaterial. Die Kategoriendefinitionen wurden nach Krippendorff (2018) induktiv, also direkt aus dem Material, in einem Verallgemeinerungsprozess, entwickelt. Es verlief nach dem empfohlenen Prozessmodell. So wurde das Thema „Unterrichtsqualität“ theoriegeleitet bestimmt (Kapitel 2) und auf diese Weise ein Selektionskriterium erstellt, um Unwesentliches auszusortieren. Ebenso wurde durch zusammenfassende Inhaltsanalyse ein Abstraktionsniveau der Kategorien erreicht. Als Beispiel wurden die Kategorien Bewusste Entscheidungen, Leh-

rerverhalten und Vermittlung von chemischem Fachwissen gebildet, welche in der Theorie teilweise unter den Begriff Klarheit der Lehrperson fallen. Eine exakte Definition von Letzterem konnte nicht gefunden werden, und so wurden diese drei Kategorien induktiv aus dem Material Unterscheidungsmerkmale zugeordnet. Nach Bearbeitung von ca. 10 Prozent der schriftlichen Reflexionen wurde eine Revision des Kategoriensystems vorgenommen. Im Anschluss wurden dem System folgend weitere Kategorien entwickelt. Es ergaben sich schließlich insgesamt 31 Kategorien zur Bestimmung der Reflexionskompetenz. Im zweiten Schritt, der Explikation, wurde sowohl auf den Textkontext (enge Kontextanalyse) als auch auf den theoretischen Hintergrund zurückgegriffen (weite Kontextanalyse). Dies beinhaltete eine Prüfung, ob die zu erklärende Textstelle noch in gleicher bzw. sehr ähnlicher Form in der weiteren Reflexion auftauchte. Zusätzlich waren Punkte der weiten Kontextanalyse klar, wie Informationen über die Entstehungsbedingungen des Textes und das theoretische Vorverständnis. Durch die Sammlung sämtlicher Textstellen und die Paraphrasierung dieser wurden die Kategorien geschärft. Die Strukturierung der Kategorien verlief deduktiv nach bestehender Theorie. So wurde nach Reflexionsbreite (neun Kategorien) und Reflexionstiefe (22 Kategorien) unterschieden. Die Strukturierungsdimensionen wurden nach der Fragestellung abgeleitet und theoretisch begründet. Es wurden die Hauptkategorien Beschreibung, Bewertung (positiv/negativ), Alternative und Verbesserungsvorschlag, jeweils durch die Unterkategorien Begründung, Abwägen und Zielformulierungen unterteilt. Zusätzlich existieren die Kategorien Perspektivenübernahme mit den Unterteilungen Schülerperspektive, Fachperspektive, Theorieperspektive und Politisch/Soziale Perspektive sowie die Kategorie Literaturbezug. Im Kodiermanual wurde nach einem dreischrittigen Verfahren vorgegangen (Haußer, Mayring & Strehmel, 1982; Ulich et al., 1985). Zunächst wurde eine Definition der Kategorie gegeben, anschließend wurden Ankerbeispiele aufgeführt und schließlich die Kodierregeln genannt und auf Abgrenzungsprobleme hingewiesen, um so eine eindeutige Zuordnung zu ermöglichen. Es folgte die Pilotierung und eine Überarbeitung des Systems. Der allgemeine Ablauf findet sich in Abbildung 34.

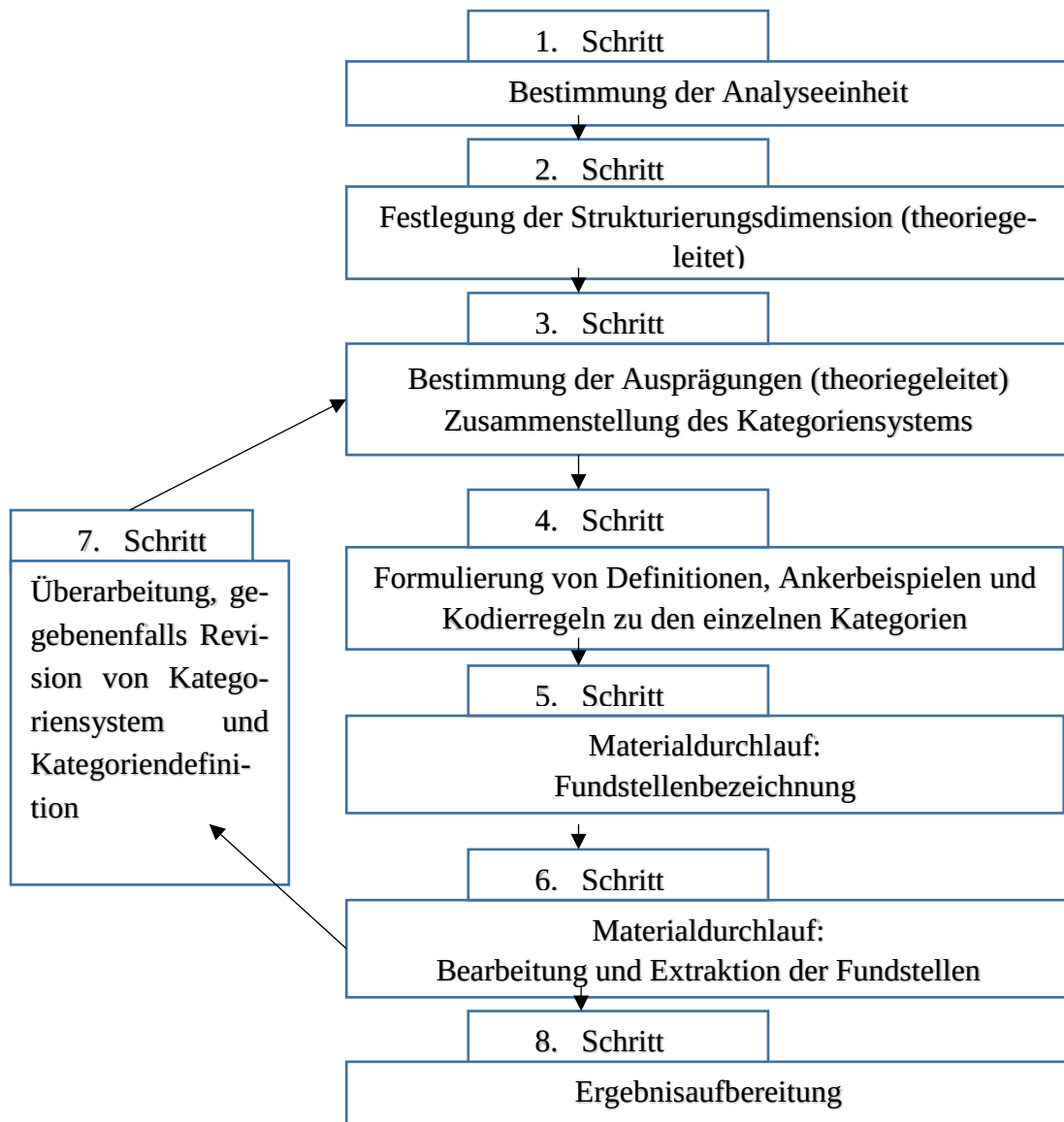


Abbildung 34: Ablaufmodell strukturierender Inhaltsanalyse (allgemein) (Mayring, 2010, S. 93)

Nach der Pilotierung:

In der Kategorie Begründungen wurden die Punkte Nennung Vorteile, Nennung Nachteile, Allgemeine Begründungen und Persönliche Meinungen zusammengefasst. Abwägen, also eine Kombination aus Vor- und Nachteilen, und Lernzielformulierungen, hier Begründungen mit inkludierten Lernzielen, wurden als Extrakategorien belassen. Grund hierfür war, dass in der Theorie zur Messung von Reflexionsprozessen durch Abwägen und Theoriebezüge oftmals ein höheres Niveau an Reflexionskompetenz erreicht werden kann, als durch alleiniges Begründen (Kapitel 3.4.1). Im Allgemeinen sind Begründungen oftmals eine Nennung von lediglich Vorteilen oder Nachteilen, weswegen eine Unterscheidung in Unterkategorien nicht möglich war.

Die Auswertung der Qualitativen Inhaltsanalyse erfolgte mithilfe der freien Software RQDA. Für die genaue Zählung der Kategoriennennungen wurde die freie Software SQLite verwendet. Hierbei wurde jeder Nennung einer Kategorie ein Punkt zugeordnet und diese im Anschluss in das Programm SPSS übertragen. So konnten unterschiedliche Entwicklungen der Reflexionsprozesse mithilfe des Mixed Methods Ansatzes deutlich

gemacht werden. Die Ergebnisse hierzu finden sich in Kapitel 6.1 und Kapitel 7, das vollständige Kodiermanual in Kapitel 10.2.2.

5.3.2. Testinstrumente und Fragebogen – Schülerinnen und Schüler

Bevor an den jeweiligen Schulen die Datenerhebung begann, musste zunächst eine schriftliche Einverständniserklärung der Schülerinnen und Schüler sowie der Erziehungsberechtigten eingeholt werden. Die Namen der Probanden wurden auf allen Fragebogen mittels Kodierungen anonymisiert. So konnte in der ganzen Studie für die Datensicherheit der Mitwirkenden gesorgt werden. Durch die Zuordnung der Codes zu den Probanden und die wiederholte Nutzung dieser in den verschiedenen Tests war eine eindeutige Zuordnung der Erhebungsinstrumente zu den jeweiligen Schülerinnen und Schülern möglich.

In der Pilotstudie war in einer Klasse der Anteil an nicht ausgefüllten Items bzw. mehrfach gekreuzten Items höher. Grund war eine zeitlich parallele Testungssituation in zwei verschiedenen Klassen aufgrund von organisatorischen Schwierigkeiten der Schule. Die eigentliche Testleiterin konnte nur in einer Klasse anwesend sein. Trotz genauer Einweisung der in der zweiten Klasse anwesenden Lehrkraft war eine passende Umsetzung der Anforderungen nicht gänzlich gewährleistet. Dies führte für die restliche Studie zum Einsatz von insgesamt zwei Testleitern, welche exakt die gleichen Arbeitsaufträge und Hinweise gaben. Ebenso wurde stets auf die Bearbeitung der Bögen in Einzelarbeit geachtet. Zusätzlich wurden die Schülerinnen und Schüler aufgefordert bei Fragen, die sie nicht beantworten konnten bzw. unsicher waren, zu raten. Dies sollte einen möglichen Vorteil von grundsätzlich ratefreudigeren Probanden ausgleichen (Bühner, 2011).

Fachwissenstest

Der Fachwissenstest wurde aus dem Projekt von Anthofer (2017) übernommen. Dieser Multiple-Choice-Test umfasste insgesamt 30 Items mit jeweils fünf ähnlich attraktiven Antwortalternativen, von denen je eine Antwort die richtige war. Im Test der Pilotstudie wurden den sechs Unterrichtsstunden fünf thematisch passende Fragen zugeordnet. In der Hauptstudie wurden aufgrund geringer Trennschärfen (< 0.15) oder nicht zufriedenstellender Itemschwierigkeit einige der Items verworfen. Ebenso wurden die Fragen zur ersten Stunde gestrichen, da diese nicht mehr gehalten wurde. Es blieben pro Unterrichtsstunde drei bis vier Fragen übrig. Um einen besseren Überblick über die Fragengruppen zu erhalten, wurde diese pro thematischer Einheit auf einer DIN A4-Seite dargestellt. Der Test wurde im Pre-Post-Follow-Up-Design durchgeführt. Die drei Messzeitpunkte fanden zwei Wochen vor dem Schülertag, am Schülertag selbst und circa zwei Monate nach dem Schülertag statt. So sind Rückschlüsse auf den Lernzuwachs möglich. Exemplarisch werden wiederum zwei Items aus den Themenblöcken Trennung von Stoffgemischen und Chromatographie aufgeführt (Abbildung 35, Abbildung 36):

Welche der folgenden Aussagen ist falsch?

Durch Sieben können Stoffe verschiedener Größe getrennt werden.	☐
Durch Dekantieren können Stoffe verschiedener Größe getrennt werden.	☒
Durch Sedimentieren können Stoffe verschiedener Dichten getrennt werden.	☐
Durch Dekantieren können Stoffe verschiedener Dichten getrennt werden.	☐
Durch Filtration kann die dekantierte Flüssigkeit zusätzlich gereinigt werden.	☐

Abbildung 35: Beispielitem aus dem Fachwissenstest zum Thema Trennung von Stoffgemischen – Schülerinnen und Schüler

Welches der folgenden Stoffgemische kann nicht durch Chromatographie getrennt werden?

Filzstiftfarbe	☐
Kaffee	☐
Salzwasser	☒
Cola	☐
Ätherische Öle	☐

Abbildung 36: Beispielitem aus dem Fachwissenstest zum Thema Chromatografie – Schülerinnen und Schüler

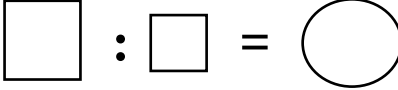
Für die Bearbeitung hatten die Schülerinnen und Schüler 20 Minuten Zeit. Pro Frage existierte eine richtige Antwortmöglichkeit, welche mit einem Punkt bewertet wurde. Bei falsch angekreuzten Aufgaben erhielten die Schülerinnen und Schüler keine Punkte. Die Gütekriterien und die in dieser Studie erzielten Ergebnisse finden sich in den Kapiteln 5.4.1, 6 und 7.

Kognitiver Fähigkeitstest

Heller und Perleth (2000) entwickelten einen Test zur Erhebung der kognitiven Fähigkeiten von Schülerinnen und Schülern der vierten bis zur zwölften Jahrgangsstufe. Dieser wurde eingesetzt, um Informationen zur Vergleichbarkeit der Klassen in dieser Studie zu erhalten und die Werte als Kontrollvariable in die Auswertung einzubeziehen. Mit dieser Zielsetzung konnten der Test auf eine Subskala beschränkt und eine Übertestung der Probanden vermieden werden. Die 25 Items begannen jeweils mit einem Paar von Figuren, welche in einer bestimmten Art und Weise zusammenpassten. Die Aufgabe bestand darin, zu entscheiden, in welcher Beziehung die gegebenen Figuren standen, und eine dritte gegebene Figur entsprechend dieser zu ergänzen. Die Probanden konnten dabei aus fünf

Figurenmöglichkeiten wählen. Zur Verdeutlichung wird exemplarisch das Erklärbeispiel aus der Originalquelle verwendet (Tabelle 32).

Tabelle 32: Beispielitem KFT-Test – Schülerinnen und Schüler (Heller & Perleth, 2000, S. 103)

B1		:	

Als erstes Paar steht ein großes Quadrat im Verhältnis zu einem kleineren Quadrat. Der große Kreis muss folglich mit dem kleineren Kreis, also Antwortmöglichkeit B, ergänzt werden. Zusätzlich wurden allgemeine Daten abgefragt, wie Alter, Geschlecht der Schülerinnen und Schüler und deren Zeugnisnote in Physik des vorangegangenen Schuljahres. Für die einmalige Testung hatten die Probanden acht Minuten Zeit zur Bearbeitung. Pro richtiger Zuordnung erhielten die Testpersonen einen Punkt.

Fragebogen zum Fachinteresse

Aus dem Handbuch PISA (Frey et al., 2009; Pekrun, 2005, S. 50) wurden wie im Projekt von Anthofer (2017) die Skala 3.1.28 zur Freude und zum Interesse an Naturwissenschaften ($N_{Items}=5$) und die Skala 3.1.60 zum naturwissenschaftlichen Fähigkeitskonzept ($N_{Items}=6$) entnommen und adaptiert. Dieser Fragebogen kam einmalig vor dem Schülertag als Kontrollvariable zum Einsatz. Dabei bewerteten die Schülerinnen und Schüler Aussagen auf einer sechsstufigen Likert-Skala von „stimme voll zu“ bis „stimme gar nicht zu“. Exemplarisch werden ausgewählte Beispielitems aufgeführt (Tabelle 33):

Tabelle 33: Beispielitem Fachinteresse – Schülerinnen und Schüler

	<i>stimme voll zu</i>	<i>stimme zu</i>	<i>stimme eher zu</i>	<i>stimme eher nicht zu</i>	<i>stimme nicht zu</i>	<i>stimme gar nicht zu</i>
Es macht mir Spaß, mich mit chemischen Themen zu beschäftigen.	☐	☐	☐	☐	☐	☐
Ich lese gerne etwas über Chemie.	☐	☐	☐	☐	☐	☐
Ich beschäftige mich gerne mit chemischen Problemen.	☐	☐	☐	☐	☐	☐

Hierbei wurde die Likert-Skala mit den Punkten eins bis sechs belegt (1 = stimme voll zu; 6 = stimme gar nicht zu) und einmalig der Mittelwert berechnet.

5.4. Datenauswertung

5.4.1. Testgütekriterien

Quantitativer Teil

Die drei Kriterien Objektivität, Reliabilität und Validität bestimmen die Güte eines Tests (Bühner, 2011). Diese drei Begriffe lassen sich weiter ausdifferenzieren.

Objektivität

Der Begriff Objektivität ist definiert als „der Grad, in dem die Ergebnisse eines Tests unabhängig vom Untersucher sind“ (Bühner, 2011, S. 58). Es lassen sich Durchführungsobjektivität, Auswertungs- und Interpretationsobjektivität unterscheiden. Erstere ist in der Regel durch ausführliche, genaue Durchführungsanweisungen gesichert (Rost, 2013). Jedem Probanden müssen die gleichen Hilfsmittel zur Verfügung stehen und Beeinflussungen sind auszuschließen (Bortz & Döring, 2007). Im Rahmen dieser Studie war die Testung der Hauptstudie auf zwei Testleiter beschränkt. Die Leiterin der Studie gab der zweiten Testperson im Vorfeld eine ausführliche mündliche und schriftliche Einweisung. Zusätzlich waren auf fast allen quantitativen Tests und Fragebogen die Hinweise zur Bearbeitung, teils sogar mehrfach, abgedruckt. Die Auswertungsobjektivität ist durch den Einsatz von Multiple-Choice-Aufgaben bzw. Auswertungsmasken gewährleistet. Durch das geschlossene Antwortformat ist eine Beeinflussung der Beurteilung der Antworten durch die Probanden ausgeschlossen (Bortz & Döring, 2007). Ebenso ist es durch die Bewertung der Tests jedem Auswerter gleich möglich, auf dieselben Punkte- oder Leistungswerte zu kommen (Bühner, 2011). Mittels der exakten Bestimmung eines Wertes kann nach Lienert und Raatz (1998) auch die Interpretationsobjektivität verwirklicht werden. Die Daten dieser Studie wurden durch die Testleiterin in SPSS eingefügt. Zusätzlich arbeitete eine neutrale Person ca. 10 Prozent der Daten (4656 Items bei den Studierenden und 5670 Items bei den Schülerinnen und Schülern) erneut ein. Im Anschluss wurde eine Übereinstimmung errechnet. Es gab drei Fälle der Nicht-Übereinstimmung, sodass von einem Eingabefehler kleiner als 0.03 Prozent ausgegangen werden kann.

Reliabilität

Die Reliabilität entspricht der Messgenauigkeit bzw. der Zuverlässigkeit eines Tests. Die Messung gibt Aufschluss, wie exakt der Fragebogen ein spezifisches Konstrukt abbildet (Kelava & Moosbrugger, 2012; Lienert & Raatz, 1998). Als Reliabilitätskoeffizient findet hier der Alphakoeffizient nach Cronbach zur Messung interner Konsistenz Anwendung. Dieser wird für alle Fragebogen und Tests der Pilot- und Hauptstudie ermittelt und bei den Ergebnissen (Kapitel 6 und 7) geschildert. Die Alpha-Werte können wie folgt interpretiert werden (Tabelle 34):

Tabelle 34: Interpretation der Alpha-Werte (George & Mallery, 2002, S. 231)

α	Interpretation
> .90	Exzellent
> .80	Gut
> .70	Akzeptabel
> .60	Fragwürdig
> .50	Schlecht
$\leq$.50	Inakzeptabel

Somit wird bei einem Wert größer .70 eine angemessene Reliabilität eines Fragebogens angenommen. Moosbrugger und Kelava (2012) folgern weiter, dass bei einem Einsatz eines Testverfahrens, für welches kein geeigneteres anderes Verfahren existent ist, ein niedrig reliables Messinstrument aufschlussreicher sein kann als ein Verzicht des Einsatzes. Aufgrund der Breite des abgetesteten Konstrukts des Wissenstests, erscheint eine Reliabilität von .72 als erwartbar zufriedenstellend. Im Folgenden werden die Reliabilitäten der Tests aus dem Projekt von Anthofer (2017) angegeben, da diese Studie darauf aufbaut und dadurch Vergleiche ermöglicht (Tabelle 35):

Tabelle 35: Ergebnisse der Tests nach Anthofer (2017)

	Test	Statistische Kennwerte (Hauptstudie)
Studierende (N = 37)	Test zum fachdidaktischen Wissen	$\alpha = .77$ (37 Paarvergleiche) Trennschärfe > 0.15
	Fachwissenstest	$\alpha = .80$ (21 Items) Trennschärfen > 0.15
	Studierendenfragebogen	$\alpha_{\text{Selbstbezogene Ziele}} = .65$ (16 Items) $\alpha_{\text{Fachspezifische Überzeugungen}} = .72$ (39 Items) $\alpha_{\text{Selbstwirksamkeitserwartung}} = .91$ (22 Items)
Schülerinnen und Schüler (N = 169)	Wissenstest	$\alpha = .72$ (30 Items) Hauptstudie Trennschärfe > 0
	Kognitiver Fähigkeitstest	$\alpha = .79$ (25 Items)
	Fachinteresse	$\alpha_{\text{Gesamt}} = .87$ (11 Items) $\alpha_{\text{Freude/Interesse an Chemie}} = .83$ (5 Items) $\alpha_{\text{Chemiebezogenes Fähigkeitsselbstkonzept}} = .82$ (6 Items)

Validität

Lienert und Raatz (1998) definieren Validität als Angabe über die Exaktheit eines Tests, sprich als Angabe, wie genau das zu messende Merkmal auch wirklich gemessen wird. Es lassen sich drei Validitätsarten unterscheiden: Inhaltsvalidität, Kriteriumsvalidität und Konstruktvalidität (Bühner, 2011).

Inhaltsvalidität:

Von Inhaltsvalidität ist die Rede, wenn ein Test und auch jedes einzelne Item das angestrebte Konstrukt präzise misst (Bühner, 2011). Hierbei soll auf Basis logischer bzw. fachlicher Überlegungen die Inhaltsvalidität bestimmt werden (Michel & Conrad, 1982). Im Rahmen der vorliegenden Arbeit wurde der Reflexionskompetenztest für die Studierenden konzipiert und auf gängiger Fachliteratur aufgebaut. Zusätzlich wurde der Test Universitätsmitarbeitern und -professoren zur Korrektur vorgelegt. Die Verbesserungsvorschläge sind alle umgesetzt.

Kriteriumsvalidität

Um die Kriteriumsvalidität zu überprüfen, erfolgte die Berechnung der Korrelationen zwischen Testvariablen und der Kriteriumsvariablen. Von einem diagnostischen Verfahren spricht man, wenn das Ergebnis einer Messung mit einem praktisch bedeutsamen Außenkriterium erwartungs- bzw. theoriekonform im Zusammenhang steht (Eid & Schmidt, 2014). Da der Pre-Test zum Thema Reflexionskompetenz zunächst für alle Studierenden ohne jegliches Vorwissen ablief, ergaben sich keine Korrelationen. Im Seminar wurden jedoch Inhalte zu den Themengebieten Reflexionskompetenz und durch die Stundenvorbereitungen indirekt Fachwissen zum Thema Stofftrennung vermittelt. Reflexionskompetenz ist, wie in Kapitel 3 aufgeführt, abhängig vom Wissen zur jeweiligen Theorie des Fachgebietes (Hatton & Smith, 1995; Leonhard & Rihm, 2011; Wyss, 2013; Zeichner & Liston, 1996; Zimmermann & Welzel, 2008). Es konnte in den Post-Tests ein mittlerer Effekt zwischen den Ergebnissen des Reflexionskompetenztests und des Fachwissens-tests erreicht werden ($r_s = .319$, $p = .047$, $n = 39$). Zur Orientierung, wie groß der gefundene Zusammenhang ist, folgt die Einteilung von Cohen (1992, S. 156):

$r = .10$ entspricht einem schwachen Effekt

$r = .30$ entspricht einem mittleren Effekt

$r = .50$ entspricht einem starken Effekt

Die Veränderung der Korrelation nach Spearman zwischen Reflexionskompetenz und Fachwissen von Pre- zu Posttest kann vorsichtig als Hinweis auf vorhandene Kriteriumsvalidität interpretiert werden.

Konstruktvalidität

Die letzte Validitätsart soll durch Prüfung von Zusammenhangsstrukturen erfolgen, welche aus formal theoretischen Annahmen abgeleitet wurden (Moosbrugger & Kelava, 2012). Die theoretischen Annahmen können dem Kapitel 3 entnommen werden. Ebenso soll es möglich sein, mit den aus dem Test erhaltenen Werten Hypothesen bestätigen zu können (Bortz & Döring, 2007), siehe Kapitel 4, 6 und 7.

Qualitativer Teil

Für die Qualitative Inhaltsanalyse existieren nach Krippendorff (1980) spezielle inhaltsanalytische Gütekriterien (Abbildung 37).

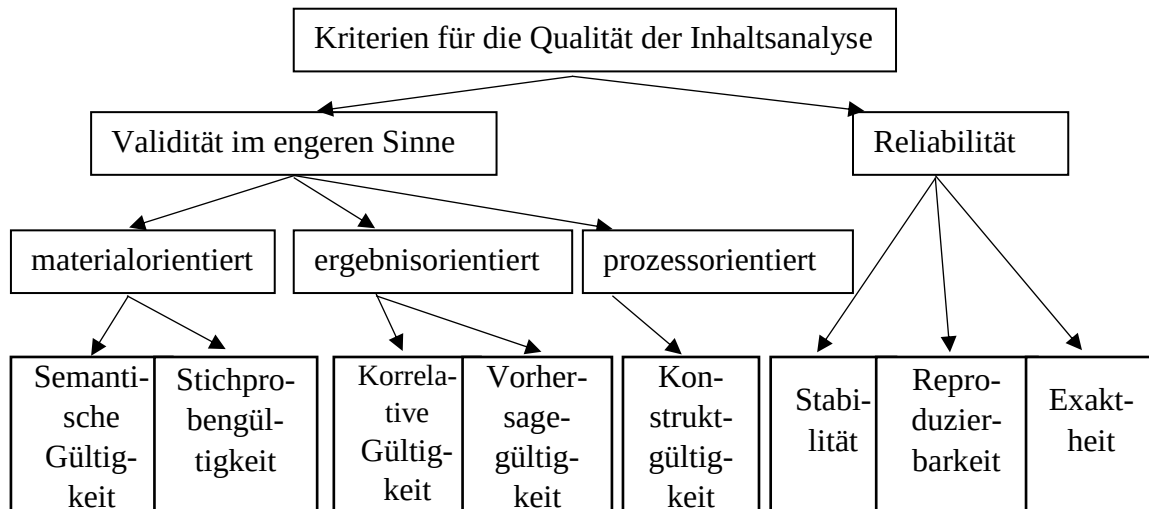


Abbildung 37: Inhaltsanalytische Gütekriterien nach Krippendorff (1980, S. 158)

Validität im engeren Sinne

Semantische Gültigkeit

Die Semantische Gültigkeit entspricht der Richtigkeit der Bedeutungsrekonstruktion, die wiederum der Angemessenheit der Kategoriendefinitionen entspricht. Sie soll aus drei Teilen – Definition, Ankerbeispiele und Kodierregeln – bestehen (Mayring, 2010). Diese Einteilung wurde im Kodiermanual eingehalten. Ebenso wurden die von Krippendorff (1980) empfohlenen Überprüfungen vorgeschlagen, welche den Vergleich aller gesammelten Textstellen zu einem Konstrukt mit Überprüfung der Homogenität beinhaltet. Dies erfolgte in dieser Studie mithilfe der freien Software RQDA. Ebenso wurden Problemfälle im Kodiermanual erörtert.

Stichprobengültigkeit

Die Stichprobe wurde von den Studierenden festgelegt, welche das Pflichtseminar Planung und Analyse von Chemieunterricht belegt haben. Somit entsprachen sich Stichprobe und Anwendungsfeld.

Korrelative Gültigkeit

Hierbei werden Korrelationen mit einem Außenkriterium überprüft. Als sinnvoll gilt der Vergleich mit Ergebnissen, die durch andere Methoden wie zum Beispiel Tests gewonnen wurden (Mayring, 2010). Es konnte eine positive Korrelation der Reflexionstiefe mit dem Konstrukt Lernen von Chemie in der Schule erzielt werden ($r_s = .265$, $p = .045$, $n = 58$). Dabei befragen die einzelnen Items die Einstellung zur Gestaltung des Unterrichts, also exakt die Punkte, über welche sich die Probanden in der Reflexionstiefe Gedanken machen. Zur besseren Veranschaulichung werden diejenigen Items aufgeführt, welche von den Probanden auf einer fünfstufigen Likert-Skala bewertet werden (Tabelle 36).

Tabelle 36: Items des Konstrukts Lernen von Chemie in der Schule

Lernen von Chemie in der Schule
1. SuS benötigen ausführliche Anleitung dazu, wie Experimente durchzuführen sind.
2. Einzelarbeit ist für den Lernerfolg im Chemieunterricht effektiver als Gruppenarbeit.
3. Lehrkräfte sollten für die Durchführung von Experimenten detaillierte Vorgehensweisen vermitteln.
4. SuS sollte häufig Gelegenheit gegeben werden, Experimente in Paaren oder Kleingruppen durchzuführen.
5. SuS sollten bereits <u>Anwendungsaufgaben</u> erhalten, bevor sie alle relevanten Konzepte und Prozeduren gut beherrschen.
6. SuS benötigen ausführliche Anleitung dazu, wie Anwendungsprobleme zu lösen sind.
7. SuS sollte häufig Gelegenheit gegeben werden, in Paaren/Kleingruppen Anwendungsprobleme zu lösen.
8. Lehrkräfte sollten bei Schwierigkeiten für das Lösen von Aufgaben <u>detaillierte Vorgehensweisen</u> vermitteln.

Ebenso konnte eine positive Korrelation der Perspektivenübernahme mit dem obigen Konstrukt Lernen von Chemie in der Schule berechnet werden ($r_s = .279, p = .034, n = 58$) sowie mit dem Konstrukt Bestrebungen im Studium ($r_s = .304, p = .020, n = 58$) (Tabelle 37).

Tabelle 37: Items des Konstrukts In meinem Studium strebe ich danach...

In meinem Studium strebe ich danach ...
1. ... neue Ideen zur Vermittlung von Wissen in meinem Fach zu bekommen.
2. ... komplizierte Vorlesungen besser zu verstehen.
3. ... meine pädagogischen Kompetenzen zu verbessern.
4. ... meine fachlichen Kompetenzen zu verbessern.
5. ... meine fachdidaktischen Kompetenzen zu verbessern.
6. ... komplexe Inhalte meines Faches wirklich zu begreifen.
7. ... mit wenig Aufwand durch das Studium zu kommen.
8. ... mich in meinem Fach zunehmend besser auszukennen.
9. ... den Prozess der Wissensvermittlung in meinem Fach wirklich zu begreifen.
10. ... meinen Kommilitonen zu zeigen, dass ich mehr weiß als andere.
11. ... bei meinen Kommilitonen nicht den Eindruck zu erwecken, dass ich mein Studium weniger gut bewältige als andere.
12. ... vor meinen Kommilitonen zu verbergen, wenn ich etwas weniger gut kann als andere.
13. ... meinem Dozenten zu beweisen, dass ich mehr weiß als andere.
14. ... vor meinem Dozenten zu verbergen, wenn ich etwas weniger gut kann als andere.
15. ... dass ich mir selbst nicht vorhalten muss, ich würde mein Studium weniger gut bewältigen als andere.
16. ... mir selbst nicht eingestehen zu müssen, wenn ich etwas weniger gut kann als andere.

Besonders die Korrelationen mit dem Konstrukt Lernen von Chemie in der Schule und dem Konstrukt Bestrebungen im Studium mit der Übernahme anderer Perspektiven zeigen eine Verlagerung des Blickwinkels der Probanden von lediglich der Betrachtung der Bedürfnisse und Ziele als Lehrperson hin zu Überlegungen über die Bedürfnisse ihrer Schülerinnen und Schüler in der Schülerperspektive oder über fachliche bzw. didaktische Aspekte in Fach- oder Theorieperspektive.

Vorhersagegültigkeit

Zur Vorhersagegültigkeit kann im Rahmen dieser Studie keine Aussage getroffen werden.

Konstruktvalidität

Die Konstruktvalidität ist überprüfbar durch den Vergleich mit ähnlichen Konstrukten und/oder Situationen bzw. durch den Vergleich mit bestehenden etablierten Theorien und Modellen und schließlich den Vergleich mit Erfahrungen aus dem Kontext des vorliegenden Materials (Mayring, 2010). Dies wurde bereits in den Kapiteln 3.4.1, und 3.4.2 beschrieben.

Reliabilität

Stabilität

Ob ein Kodiermanual stabil ist, kann durch mehrfache Anwendung des Instruments auf das Material überprüft werden. Dies wird auch als Intracoderreliabilität bezeichnet (Mayring, 2010). Hierbei werden, wie auch bei der Interdecoderreliabilität, drei Koeffizienten angegeben. Die prozentualen Übereinstimmungen ermittelt man aus Kreuztabellen in SPSS. Der Wert zieht zur Berechnung des Durchschnitts die Summe aller gleich kodierten Paare heran, welche durch die Summe aller, sprich gleich und ungleich, kodierten Paarungen dividiert wird. So folgt eine Unterschätzung der eigentlichen Übereinstimmung, da einseitige Nichtkodierungen als ungleiches Paar und somit doppelt gewichtet werden. Dieser Nachteil wird durch den als einfachstes Reliabilitätsmaß bezeichneten Holsti-Koeffizienten nivelliert, welcher ebenso aus Kreuztabellen in SPSS ermittelt werden kann (Früh, 2015; Mayring, 2010) (Abbildung 38):

$$\frac{(\text{Zahl der Kodierer}) \times (\text{Zahl der übereinstimmenden Urteile})}{(\text{Zahl aller Kodierurteile})}$$

Abbildung 38: Formel zur Berechnung der Intra-/Interdecoderreliabilität nach Holsti (Holsti, 1969, S. 140)

Es gelten folgende Interpretationen (Tabelle 38):

Tabelle 38: Interpretation des Holsti-Koeffizienten

Holsti-Koeffizient	Interpretation bei hoch-inferenten Kodierungen
$c_r > .75$	Sehr gute Übereinstimmung
$.75 > c_r > .60$	Gute Übereinstimmung
$.60 > c_r > .40$	Akzeptable Übereinstimmung
$c_r < .40$	Unzureichende Übereinstimmung

Der Nachteil dieses Koeffizienten ist die fehlende zufällig zu erwartende Übereinstimmung (Mayring, 2010). Dies ist von Bedeutung, wenn nur aus wenigen Kategorien ausgewählt werden kann. Je mehr Kategorien ein Kodiermanual hat, umso unwahrscheinlicher ist eine zufällig gleiche Kodierung. Da das Kodiermanual in dieser Studie in der Reflexionsbreite zehn Kategorien beinhaltet, ist es also für diesen Wert wichtig die zufällige Übereinstimmung herauszurechnen. Um diesem Problem Rechnung zu tragen, empfiehlt Krippendorff (1980) seinen Koeffizienten. Dieser inkludiert beobachtete Nicht-Übereinstimmungen und zufällig zu erwartende Übereinstimmung ebenso wie eine mögliche Berechnung für mehrere Kodierer (Wirtz & Caspar, 2002). Für Krippendorffs α ist es möglich, ein kostenloses Makro von Hayes und Krippendorff (2007) für SPSS zu installieren¹. Somit war ebenfalls eine Berechnung dieses Koeffizienten möglich (Ergebnisse in Kapitel 5.4.9). Aufgrund der aufgeführten Vor- und Nachteile der Koeffizienten werden in dieser Studie zunächst neben der prozentualen Übereinstimmung beide Koeffizienten berechnet und angegeben. Da jedoch nur geringe Unterschiede zu sehen sind (Kapitel 6.1), erfolgt im Weiteren die Angabe der prozentualen Übereinstimmung ergänzt durch Krippendorffs Alpha.

Reproduzierbarkeit

Mayring (2010, S. 120) definiert diesen Begriff als „Grad, in dem die Analyse unter anderen Umständen, anderen Analytikern zu denselben Ergebnissen führt.“ Gemessen wird dies mithilfe der InterCoderreliabilität. Neben der Leiterin dieser Studie wurden insgesamt 20 Prozent der Messzeitpunkte von zwei weiteren Kodierern analysiert. Hierbei werden erneut, wie im oberen Absatz zur Stabilität erklärt, die drei Koeffizienten prozentuale Übereinstimmung, Holsti-Koeffizient und Krippendorffs α angegeben (Ergebnisse in Kapitel 6.1.5).

Exaktheit

Für dieses Reliabilitätsmaß wurden die vier Quellen der Nicht-Reliabilität nach Krippendorff (1980) überprüft.

1. Auswertungseinheit: Es wurde überprüft, ob Unstimmigkeiten in der Auswertungseinheit bei mehreren Kodierern systematisch erfolgen.
2. Analytiker: Die InterCoderreliabilität wurde untersucht.
3. Einzelne Kategorien: Es wurde überprüft, ob Unstimmigkeiten bei bestimmten Kategorien besonders häufig vorkommen. Dies wurde durch klare Definitionen behoben.
4. Kategoriendifferenzierung: Die Kategoriendifferenzierung wurde überprüft und unscharfe Kategorien zu größeren Kategorien zusammengefügt.

5.4.2. Wissenstests

Im Rahmen dieser Studie durchliefen die Probanden, angelehnt an Anthofer (2017), insgesamt drei Wissenstests: den Fachwissenstest, den fachdidaktischen Test und den Reflexionskompetenztest. Die Schülerinnen und Schüler wurden sowohl auf das Fachwissen wie auch auf ihre kognitiven Fähigkeiten getestet.

¹ <http://afhayes.com/index.html>

Nach Eingabe der Rohdaten in das Computerprogramm SPSS wurden richtige Antworten mit einem Punkt bewertet und falsche Antworten mit null Punkten. Die erreichte Punktzahl wurde im Anschluss durch die Gesamtpunktzahl dividiert und mit 100 multipliziert. Diese Prozentangaben werden als Scores bezeichnet und bieten die Möglichkeit eines direkten Vergleichs unterschiedlich umfangreicher Tests. Fast alle Wissenstests bestehen aus Multiple-Choice-Aufgaben im Single-Select-Format.

Eine Ausnahme stellt der fachdidaktische Test dar, wobei die Studierenden zu einem Handlungsstamm vier mögliche Alternativen mit jeweils einer sechsstufigen Likert-Skala zur Auswahl hatten. Daraus wurden Paarvergleiche gezogen, welche durch ein Experten-rating zuvor validiert wurden. Stimmte die Richtung von Paarvergleich und Experten-rating überein, wurde ein Punkt vergeben. Unter „Richtung“ ist zu verstehen, welches der jeweils zwei Items eines Paarvergleichs besser als das andere geratet wurde. Es wurden nur die Paarvergleiche in die Bepunktung einbezogen, die sich in früheren Studien bewährt hatten. Final wurden 38 Paarvergleiche bei 19 Items in die Auswertung der Hauptstudie einbezogen (Anthofer, 2017; Backes et al., 2012).

5.4.3. Meinungsfragebogen

Der Studierendenfragebogen setzte sich zunächst aus persönlichen Angaben zusammen, wie Alter, Geschlecht, Abiturschnitt und Fragen zur bisherigen Ausbildung. Im Anschluss wurden die Konstrukte Zielorientierung, fachspezifische Überzeugungen und Selbstwirksamkeitserwartungen erfragt. Obwohl hierunter z. T. stabile Merkmale wie das Fachinteresse fielen, wurde auch dieser Fragebogen im Prä-Post-Follow-Up Design erhoben, um die Testzeit konstant zu halten und an allen Messzeitpunkten vergleichbar vorzugehen. Auf Seiten der Schülerinnen und Schüler fand eine einmalige Erhebung des Fachinteresses statt, da hier aufgrund nur eines Schülertages keine Änderungen zu erwarten sind.

Die Likert-Skala wurde dabei mit den Zahlen eins bis fünf belegt und entsprechend den Antworten der Probanden in SPSS eingetragen. Zuvor wurden sämtliche negativ formulierte Items umkodiert (Bühl, 2014). Zu Vergleichszwecken wurde das arithmetische Mittel gebildet. Da zunächst eine ordinale Skalierung vorlag, wurde zusätzlich der exakte Median mithilfe von SPSS berechnet (Clauss, Finze & Partzsch, 2004).

5.4.4. Parametrische und nichtparametrische Tests

Die inferenzstatistischen Methoden können in die Bereiche parametrische (verteilungsabhängige) und nichtparametrische (verteilungsunabhängige) Analysen eingeteilt werden (Rost, 2013). Zur fehlerfreien Anwendung müssen bestimmte Voraussetzungen erfüllt sein, wie beispielsweise eine bestimmte Stichprobengröße und/oder Skalenqualitäten von mindestens Intervallskalierung, und/oder Verteilungsannahmen wie Normalverteilung und eine im Falle von drei Messzeitpunkten vorliegende Sphärizität (Field, 2013). Nichtparametrische Verfahren verlangen weniger Voraussetzungen an die Daten und sind vor allem bei kleinen Stichproben anzuwenden (Rost, 2013). Abhängig von der Anzahl der zu vergleichenden Gruppen und ob das Stichprobenverhältnis abhängig oder unabhängig ist, finden unterschiedliche parametrische und nichtparametrische Verfahren Anwendung.

Varianzhomogenität

Für die Anwendung parametrischer Verfahren ist es Voraussetzung, dass in beiden Populationen die Varianzen eines Merkmals identisch sind (Bortz & Schuster, 2011). Zur Überprüfung eignen sich sowohl der F-Test als auch der Levene-Test. Da ersterer sich jedoch gegenüber einer Verletzung der Normalverteilung nicht robust verhält, wird in dieser Studie der Levene-Test bevorzugt (Bortz & Schuster, 2011). Als Nullhypothese wird angenommen, dass sich die Varianzen nicht unterscheiden. Ist hierbei die Irrtumswahrscheinlichkeit größer .05, wird die Nullhypothese beibehalten. Man geht also davon aus, dass sich die Varianzen nicht signifikant unterscheiden (Bühl, 2014).

Normalverteilung

Eine weitere Voraussetzung zur Anwendung parametrischer Methoden ist die Annahme der Normalverteilung eines Merkmals in der Gesamtpopulation (Raab-Steiner & Benesch, 2010). Ist keine Normalverteilung gegeben, werden die entsprechenden nichtparametrischen Verfahren angewandt. Die Überprüfung erfolgt mithilfe des Kolmogorov-Smirnov-Tests. Ein signifikanter Unterschied wird ab einer Irrtumswahrscheinlichkeit von $p < .05$ angenommen (Bühl, 2014). In diesem Fall geht man von keiner Normalverteilung aus.

Sphärizität

Wenn die Varianzen der Differenzen zwischen den einzelnen Messzeitpunkten homogen sind, spricht man von Sphärizität. Ist diese gegeben kann eine einfaktorielle Varianzanalyse berechnet werden. Zur Überprüfung wird der Mauchly-Test gerechnet. Fällt dabei das Testergebnis signifikant aus, muss von einer fehlenden Sphärizität ausgegangen werden (Field, 2013).

Mittelwertvergleiche

Bei den aufgeführten Tests werden, je nachdem ob parametrisch oder nichtparametrisch, die Mittelwerte bzw. die Mediane verschiedener Stichproben miteinander verglichen. Die Testung auf signifikante Unterschiede kann nach Bühl (2014, S. 177) wie folgt interpretiert werden (Tabelle 39):

Tabelle 39: Werte der Signifikanz nach Bühl (2014, S. 177)

<i>p</i>	Bedeutung	Symbol
$> .05$	Nicht signifikant	n. s.
$\leq .05$	Signifikant	*
$\leq .01$	Sehr signifikant	**
$\leq .001$	Höchst signifikant	***

5.4.5. Residuenberechnung

Residuen-Berechnungen erlauben einen genaueren Blick auf die Nachtestergebnisse, da diese in Abhängigkeit der Vortestergebnisse dargestellt werden. Residuen lassen sich als Abweichungen vom Erwartungswert (z. B. einer Regressionsgeraden) interpretieren. Unterschiede der Residuen können ggf. auf die Zugehörigkeit zu verschiedenen Gruppen

wie z. B. Interventions- und Kontrollgruppe zurückgeführt werden. Weiter können Residuenanalysen darüber informieren, ob die Voraussetzungen für inferenzstatistische Auswertungen, wie etwa Signifikanztests, gegeben sind (Bortz & Schuster, 2011).

5.4.6. Effektstärke

Werden statistisch signifikante Unterschiede zwischen Gruppen gefunden, ist es sinnvoll, nach der Relevanz bzw. der praktischen Bedeutsamkeit zu fragen (Rost, 2013). Eine statistisch signifikante Differenz ist von mehreren Faktoren abhängig, wie der in der Stichprobe vorliegenden Größe des Effekts, dem gewählten Signifikanzniveau und der Stichprobengröße. Das Effektmaß d ist bei unabhängigen Stichproben definiert als die „Differenz der Mittelwerte beider Gruppen (M_A minus M_B), geteilt durch die gepoolte Standardabweichung“ (Rost, 2013, S. 240). Bühner (2011, S. 268) teilt die Effektstärke d wie folgt ein (Tabelle 40):

Tabelle 40: Interpretation der Effektstärke d

d	Interpretation
$> .20$	Kleiner Effekt
$> .50$	Mittlerer Effekt
$> .80$	Großer Effekt

Gleiche Einteilungen finden sich bei Cohen (1988, S. 24–27), welche auch von Bortz und Schuster (2011, S. 109) vorgeschlagen werden.

Ein weiteres Maß für die Effektstärke wäre r , welches nach Cohen (1992) wie folgt eingeteilt wird (Tabelle 41):

Tabelle 41: Interpretation der Effektstärke r

r	Interpretation
$> .10$	Schwacher Effekt
$> .30$	Mittlerer Effekt
$> .50$	Großer Effekt

Bei einfaktoriellen Varianzanalysen wird das Effektstärkemaß η^2 (η^2) angewandt (Bortz & Döring, 2007). Dieses Maß gibt den „Anteil der Varianz der abhängigen Variablen an, der durch die unabhängige Variable erklärt wird“ (Janssen & Laatz, 2013, S. 268). Bühner (2011, S. 268) teilt die Effektstärke η^2 wie folgt ein (Tabelle 42):

Tabelle 42: Interpretation der Effektstärke η^2

η^2	Interpretation
0.01	Kleiner Effekt
0.06	Mittlerer Effekt
0.14	Großer Effekt

5.4.7. Korrelationen

In dieser Studie wurden zur Testung der Validität bereits Spearman Korrelationen für nonparametrische Daten berechnet. Dabei wurden die Einteilungen nach Cohen (1992, S. 156) übernommen. Ebenso wurden Korrelationskoeffizienten nach Pearson berechnet,

um Zusammenhänge zwischen zwei Variablen als Produkt-Moment-Korrelation zu erhalten. Hier können die Interpretationen von Bühner und Ziegler (2009, S. 177) übernommen werden (Tabelle 43):

Tabelle 43: Unterscheidung der Korrelationen nach Pearson und Spearman

Korrelation	Effektgröße	Klassifikation		
		Schwach	Mittel	Stark
Pearson	r_p	.10	.30	.50
Spearman	r_s	.10	.30	.50

5.4.8. Regressionsanalysen

Eine multiple Regressionsgleichung ermöglicht eine Vorhersage einer Kriteriumsvariablen basierend auf mehreren Prädiktoren (Bortz & Schuster, 2011). Zusätzlich besteht die Möglichkeit, den Einfluss einer Drittvariablen zu kontrollieren. Die Güte des Regressionsmodells wird mithilfe des Statistikprogramms SPSS ermittelt, welches das Bestimmtheitsmaß R^2 berechnet. Die Ergebnisse dieser Variablen liegen zwischen 0 und 1 und geben Auskunft darüber, wie gut sich die Werte der abhängigen Variablen aufgrund der Werte der unabhängigen Variablen berechnen lassen. R^2 kann auch als prozentualer Anteil interpretiert werden.

Ein R^2 von 1 bedeutet, dass sich die Werte der abhängigen Variablen ausschließlich aus den Werten der unabhängigen Variablen ermitteln lassen. Ein R^2 von 0 lässt dementsprechend keine Rückschlüsse der unabhängigen Variablen auf die abhängige zu. Zusätzlich wird im Ergebnisteil die Prüfgröße F und die Signifikanz des Gesamtmodells angegeben. So ist es möglich, Aussagen über eine allgemeine oder lediglich beschränkte Gültigkeit der Regressionsrechnung zu machen (Bortz & Schuster, 2011; Bühner & Ziegler, 2009).

5.4.9. Kodiermanual

Das Kodiermanual kam zur Auswertung der schriftlichen Reflexionen an drei Messzeitpunkten der Studie zum Einsatz. Hierbei wurden die Reflexionen in zwei große Bereiche eingeteilt: Reflexionsbreite und Reflexionstiefe. Zunächst wurde die Reflexionsbreite kodiert, welche aus zehn Kategorien bestand. Diese wurden auf Basis der Pilotstudie induktiv und deduktiv ermittelt (Kapitel 6.1). Im Kodiermanual fanden sowohl die Regeln der qualitativen Inhaltsanalyse wie auch die Regeln zur Optimierung des Aufbaus eines qualitativen Kodiermanuals Beachtung.

Zunächst folgt ein Überblick über die Kategorien:

1. Reflexionsbreite

- Lehrerperformanz
- Lehrerverhalten
- Vermittlung von chemischem Fachwissen
- Regeln im Chemieunterricht
- Schüleraktivierung
- Sozialform
- Schülerexperiment
- Organisationsform
- Medien
- Sonstiges

2. Reflexionstiefe

- a. Beschreibung
 - i. Rechtfertigung
 - ii. Abwägung
 - iii. Zielformulierung
- b. Bewertung (positiv/negativ)
 - i. Rechtfertigung
 - ii. Abwägung
 - iii. Zielformulierung
- c. Alternative
 - i. Rechtfertigung
 - ii. Abwägung
 - iii. Zielformulierung
- d. Verbesserungsvorschlag/Konsequenz
 - i. Rechtfertigung
 - ii. Abwägung
 - iii. Zielformulierung
- e. Perspektive
 - i. Schülerperspektive
 - ii. Fachperspektive
 - iii. Theorieperspektive
 - iv. Politisch/Soziale Perspektive
- f. Literaturbezug

Im Anschluss wird der Aufbau des Manuals erklärt. Ebenso werden eventuell unbekannte Fachbegriffe wie Kodiereinheit oder Kontexteinheit definiert. Die zu kodierenden Kategorien sind grundsätzlich einheitlich aufgebaut, um einen besseren Überblick zu gewährleisten:

- 1. Definition der Kategorien
 - a. Allgemein
 - b. Differenzierung
- 2. Ankerbeispiele
 - a. Allgemein
 - b. Differenzierung (optional)
- 3. Kodierregeln
 - a. Allgemein
 - b. Differenzierung (optional)

Es folgen allgemeine Kodierregeln. So wird ein Messzeitpunkt beispielsweise dreimal kodiert. Beginnend mit der Reflexionsbreite werden im zweiten Durchgang die Punkte a) bis d) der Reflexionstiefe kodiert. Der dritte Durchgang kodiert die Perspektiven und den Literaturbezug. Dieses Vorgehen gewährleistet eine exaktere Kodierung, da sich eine Kodierung aller Kategorien bei einmaligem Durchgehen als fehleranfällig erwiesen bzw. stark zum Übersehen einiger Kategorien geführt hat. Die angegebenen Differenzierungen dienen einer stärkeren Abgrenzung der Kategorien. So sind die Ankerzitate positive Beispiele für die jeweilige Kategorie. Die Differenzierung begründet jeweils, warum diese Beispiele nicht der jeweiligen Kategorie zugeordnet werden sollen, stellen somit negative Beispiele dar.

In der Pilotstudie erfolgte eine zusätzliche Auswertung mit Zahlencodes für die jeweilige Kategorie, welche an den Rand geschrieben wurden. Dieses System wurde durch den Einsatz des Programms RQDA wesentlich erleichtert und optimiert, da es hierbei möglich ist, die gefundenen Textstellen mit dem jeweiligen Code farbig zu markieren. Auch Mehrfachkodierungen der gleichen Textstelle gestalteten sich so übersichtlicher. Zudem können alle Kodierungen einer Kategorie angezeigt werden. Um Fehler bei der Übertragung der Punkte in SPSS zu vermeiden, wurde die RQDA-Datei mithilfe der Software SQLite ausgewertet. Dieses Programm gibt die vorgenommenen Kodierungen pro Kodierer, je Messzeitpunkt und Kategorie, an. Ebenso lässt sich die Anzahl gelöschter Vorgänge ermitteln. Diese Ergebnisse wurden dann in SPSS übertragen.

Die Reflexionsbreite wurde dichotom bepunktet. So erfolgte eine Bewertung von einmaligen und auch mehrmaligen Nennungen eines Breitenpunkts mit 1. Wurde ein Breitenpunkt nicht genannt, wurde dies mit 0 gewertet.

Die einmalige Nennung einer Reflexionstiefenkategorie wurde mit einem Punkt bewertet. Dabei wurden aber mehrmalige Nennungen ebenfalls miteinberechnet. So entspricht die gegebene Punktzahl der Häufigkeit des Auftretens einer Kategorie. Wurde also die Kategorie Schülerperspektive dreimalig gefunden, so entsprach dies auch drei Punkten. Eine Ausnahme bildete die Kategorie Beschreibung. Da die alleinige Beschreibung einer Stunde keiner Reflexion im theoretischen Rahmen entspricht, wurde diese Kategorie in die Gesamtpunktzahl nicht miteinberechnet. Das komplette Kodiermanual ist in Kapitel 10.2.2 zu finden. Ergebnisse der Auswertungen werden im folgenden Kapitel berichtet.

6. Ergebnisse der Pilotstudie

An der Pilotstudie im Sommersemester 2017 nahmen insgesamt 18 Studierende, also sechs pro Interventionsgruppe und sechs in der Kontrollgruppe, teil. Diese unterrichteten insgesamt 107 Schülerinnen und Schüler an sechs Schülertagen. Aufgrund der kleinen Stichprobe sind die im Folgenden vorgestellten Ergebnisse lediglich als Richtwerte zu betrachten und werden nur deskriptiv berichtet. Auf ihrer Grundlage wurden Testinstrumente für die Hauptstudie angepasst. Die daraus resultierenden Ergebnisse finden sich in Kapitel 7. Alle deskriptiven Berechnungen werden in diesem Kapitel zunächst auf Basis der Gesamtstichprobe der Pilotstudie mit 18 Studierenden angegeben, es folgen deskriptiv die Ergebnisse der einzelnen Gruppen. Die Reliabilitäten werden aufgrund der kleinen Stichprobe der Pilotstudie und der lediglich deskriptiven Ergebnisdarstellung bei bestehenden Tests auf Basis der Pilotstudie in Kombination mit der Hauptstudie berechnet, um die Gültigkeit späterer Vergleiche der Ergebnisse gewährleisten zu können.

6.1. Testinstrumente und Fragebogen – Studierende

6.1.1. Test zum fachdidaktischen Wissen

In diesem Test gab es in der Pilot- und Hauptstudie 38 Paarvergleiche. Die Voranalysen wurden auf Basis der Post-Tests durchgeführt. So wurden bei der Überprüfung der Trennschärfen alle Relationen mit Werten unter .20 entfernt. Das Cronbachs Alpha verbesserte sich dadurch von einem ursprünglichen Wert von .80 (bei 67 Paarvergleichen; 17 Aufgaben mit Paarvergleichen und zwei Theoriefragen mit abgeänderter Bepunktung (vgl. Kapitel 5.3.1); maximal zu erreichende Punkte: 71) auf einen Wert von .85 (38 Relationen; maximal zu erreichende Punkte: 38). Im Folgenden sind die erreichten Punktzahlen der Gesamtheit der Pilotstudie und der drei untersuchten Gruppen deskriptiv dargestellt (Abbildung 39, Tabelle 44).

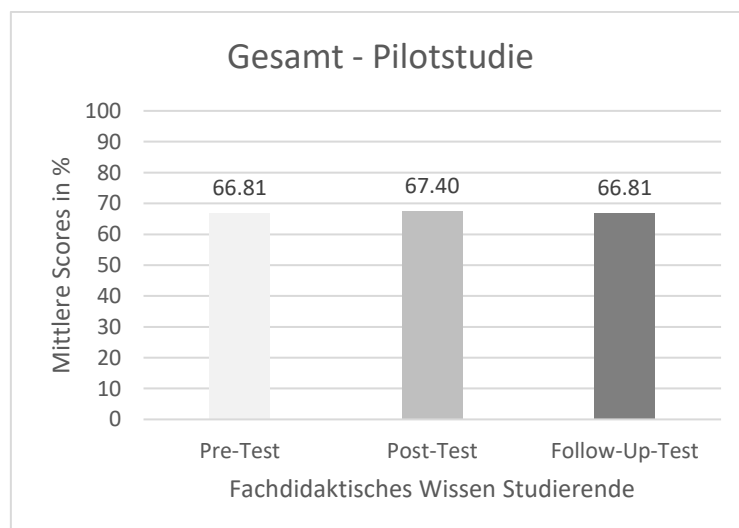


Abbildung 39: Mittlere Scores des Tests zum fachdidaktischen Wissen über drei Messzeitpunkte – Pilotstudie gesamt

Tabelle 44: Deskriptive Statistik zum fachdidaktischen Wissen über drei Messzeitpunkte – Scores Pilotstudie gesamt

Pilotstudie gesamt	N	M	SD	Max.	Min.
Pre-Test	18	66.81	13.48	84.21	44.74
Post-Test	18	67.40	19.04	89.47	21.05
Follow-Up-Test	18	66.81	14.19	94.74	39.47

Nachfolgend sind die erzielten Scores im fachdidaktischen Wissen zur Pilotstudie getrennt nach Gruppen aufgeführt (Abbildung 40, Tabelle 45, Tabelle 46, Tabelle 47).

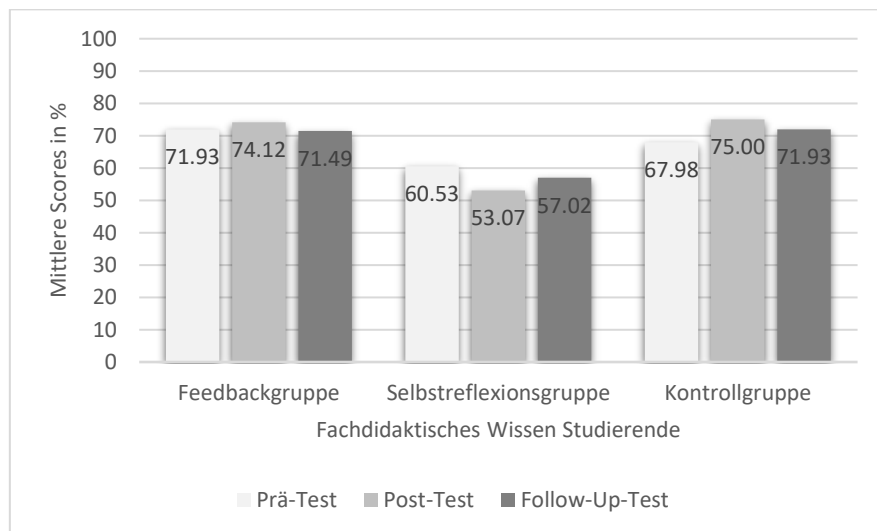


Abbildung 40: Mittlere Scores des Tests zum fachdidaktischen Wissen über drei Messzeitpunkte – Pilotstudie aufgeteilt in Gruppen

Tabelle 45: Deskriptive Statistik zum fachdidaktischen Wissen über drei Messzeitpunkte – Feedbackgruppe – Scores Pilotstudie

Feedbackgruppe	N	M	SD	Max.	Min.
Pre-Test	6	71.93	9.66	84.21	55.26
Post-Test	6	74.12	20.21	89.47	34.21
Follow-Up-Test	6	71.49	10.04	78.95	52.63

Tabelle 46: Deskriptive Statistik zum fachdidaktischen Wissen über drei Messzeitpunkte – Selbstreflexionsgruppe – Scores Pilotstudie

Selbstreflexions- gruppe	N	M	SD	Max.	Min.
Pre-Test	6	60.53	16.14	84.21	44.74
Post-Test	6	53.07	16.84	65.79	21.05
Follow-Up-Test	6	57.02	13.59	76.32	39.47

Tabelle 47: Deskriptive Statistik zum fachdidaktischen Wissen über drei Messzeitpunkte – Kontrollgruppe – Scores Pilotstudie

Kontrollgruppe	N	M	SD	Max.	Min.
Pre-Test	6	67.98	13.56	84.21	47.37
Post-Test	6	75.00	13.08	86.84	50.00
Follow-Up-Test	6	71.93	15.04	94.74	52.63

Die deskriptive Auswertung liefert erste Hinweise bezüglich einer besseren Lernwirksamkeit im Bereich Fachdidaktik für die Feedbackgruppe und die Kontrollgruppe des Seminars. Die Selbstreflexionsgruppe verschlechtert sich im Post-Test und kann diesen Lernrückschritt auch im Follow-Up-Test nicht gänzlich nivellieren. Genauere und belastbare Ergebnisse liefert die Hauptstudie.

6.1.2. Fachwissenstest

Infolge einer niedrigen Reliabilität von .69 (21 Items) wurden nach der Pilotstudie die Items 7 und 13 aufgrund negativer Trennschärfen aus dem CK Test entfernt. Ein Grund für die niedrige Reliabilität könnte ein leicht veränderter Seminarablauf im Vergleich zu Anthofers (2017) Studie sein. So wurden in dieser Studie zum Beispiel die Klassen von nur einem Studierenden unterrichtet. In der ersten Studie wurden Lehrertandems eingesetzt. Insbesondere verlagerte sich in dieser Studie der Schwerpunkt von der Entwicklung eines Seminars zur Förderung des fachspezifischen Professionswissens bei Anthofer (2017) hin zu einer Förderung der Reflexionskompetenz. Durch die Anpassung des Tests an das Studiendesign konnte, bezogen auf die Pilotstudie, ein Cronbachs Alpha von .73 (19 Items) erreicht werden (Abbildung 41, Tabelle 48).

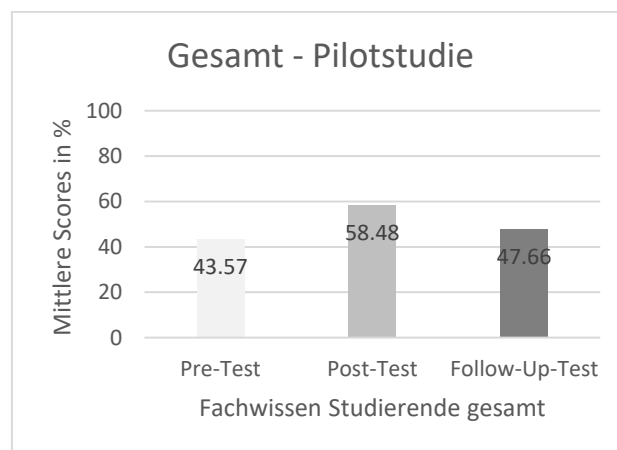


Abbildung 41: Mittlere Scores des Tests zum Fachwissen über drei Messzeitpunkte – Pilotstudie gesamt

Tabelle 48: Deskriptive Statistik zum Fachwissen über drei Messzeitpunkte – Scores Pilotstudie gesamt

Pilotstudie gesamt	N	M	SD	Max.	Min.
Pre-Test	18	43.57	15.51	63.16	0.00
Post-Test	18	58.48	16.92	94.74	31.58
Follow-Up-Test	18	47.66	14.83	84.21	21.05

Nachfolgend sind die erzielten Scores des Tests zum Fachwissen zur Pilotstudie getrennt nach Gruppen aufgeführt (Abbildung 42, Tabelle 49, Tabelle 50, Tabelle 51):

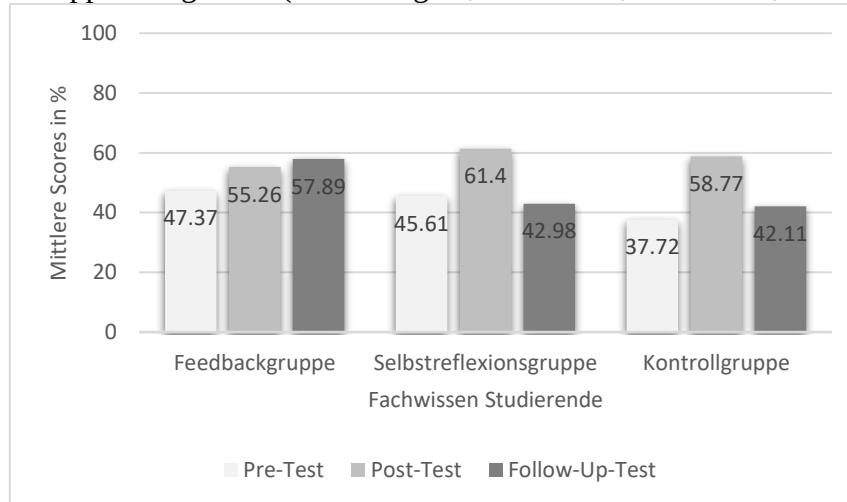


Abbildung 42: Mittlere Scores des Tests zum Fachwissen über drei Messzeitpunkte – Pilotstudie aufgeteilt in Gruppen

Tabelle 49: Deskriptive Statistik zum Fachwissen über drei Messzeitpunkte – Feedbackgruppe – Scores Pilotstudie

Feedbackgruppe	N	M	SD	Max.	Min.
Pre-Test	6	47.37	13.31	63.16	26.32
Post-Test	6	55.26	18.46	78.95	31.58
Follow-Up-Test	6	57.89	18.83	84.21	31.58

Tabelle 50: Deskriptive Statistik zum Fachwissen über drei Messzeitpunkte – Selbstreflexionsgruppe – Scores Pilotstudie

Selbstreflexionsgruppe	N	M	SD	Max.	Min.
Pre-Test	6	45.61	11.85	57.89	26.32
Post-Test	6	61.40	11.37	78.95	47.37
Follow-Up-Test	6	42.98	7.75	52.63	31.58

Tabelle 51: Deskriptive Statistik zum Fachwissen über drei Messzeitpunkte – Kontrollgruppe – Scores Pilotstudie

Kontrollgruppe	N	M	SD	Max.	Min.
Pre-Test	6	37.72	20.90	63.16	0.00
Post-Test	6	58.77	21.93	94.74	36.84
Follow-Up-Test	6	42.11	12.00	52.63	21.05

Die deskriptive Auswertung liefert erste Hinweise bezüglich einer guten Lernwirksamkeit von Pre-Test auf Post-Test im Bereich Fachwissen für alle Gruppen. Am effektivsten erscheint hierbei die Kontrollgruppe zu sein.

6.1.3. Reflexionskompetenztest

Aufgrund einer starken Itemreduzierung beim Reflexionskompetenztest zugunsten einer Trennschärfe über 0.20 konnte das Cronbachs Alpha von einem anfänglichen Wert von .60 bei 21 Items auf einen Wert von .81 mit den restlichen 12 Items gesteigert werden. Die Items 12, 15, 17 und 18 wurden für die Hauptstudie aus dem Test entfernt. Die deskriptiven Analysen führten zu folgenden Werten (Abbildung 43, Tabelle 52):

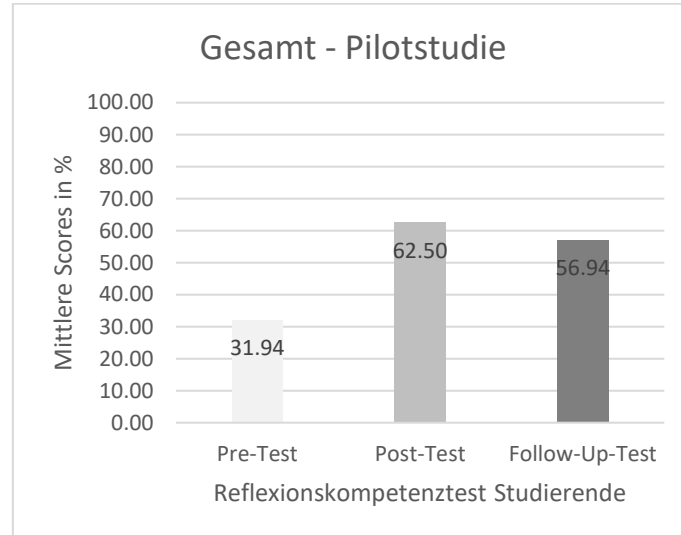


Abbildung 43: Mittlere Scores des Reflexionskompetenztests über drei Messzeitpunkte – Pilotstudie gesamt

Tabelle 52: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Pilotstudie gesamt

Pilotstudie gesamt	N	M	SD	Max.	Min.
Pre-Test	18	31.94	9.59	50.00	16.67
Post-Test	18	62.50	22.55	100.00	8.33
Follow-Up-Test	18	56.94	19.44	83.33	25.00

Aufteilung in Gruppen (Abbildung 44, Tabelle 53, Tabelle 54, Tabelle 55):

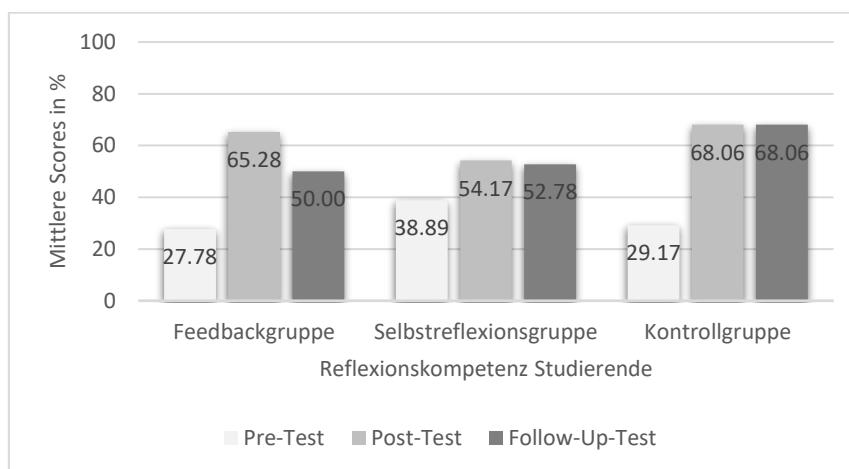


Abbildung 44: Mittlere Scores zum Reflexionskompetenztest über drei Messzeitpunkte – Pilotstudie aufgeteilt in Gruppen

Tabelle 53: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Feedbackgruppe – Scores Pilotstudie

Feedbackgruppe	N	M	SD	Max.	Min.
Pre-Test	6	27.78	4.30	33.33	25.00
Post-Test	6	65.28	14.35	83.33	50.00
Follow-Up-Test	6	50.00	22.97	83.33	25.00

Tabelle 54: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Selbstreflexionsgruppe – Scores Pilotstudie

Selbstreflexionsgruppe	N	M	SD	Max.	Min.
Pre-Test	6	38.89	11.39	50.00	16.67
Post-Test	6	54.17	20.24	91.67	33.33
Follow-Up-Test	6	52.78	13.61	75.00	33.33

Tabelle 55: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Kontrollgruppe – Scores Pilotstudie

Kontrollgruppe	N	M	SD	Max.	Min.
Pre-Test	6	29.17	8.74	41.67	16.67
Post-Test	6	68.06	31.37	100.00	8.33
Follow-Up-Test	6	68.06	18.57	83.33	33.33

Hinsichtlich des Reflexionskompetenztests können deskriptiv alle Gruppen einen Lernzuwachs verzeichnen. Feedbackgruppe und Kontrollgruppe scheinen jedoch mehr zu profitieren als die Selbstreflexionsgruppe.

6.1.4. Studierendenfragebogen

Studierendenfragebogen – Allgemeine persönliche Angaben

Der erste Teil des Studierendenfragebogens erfasste allgemeine persönliche Angaben (Tabelle 56).

Tabelle 56: Persönliche Angaben der Studierenden

Alter	M = 23,83; SD = 4,22; Max. = 37; Min. = 20
Geschlecht	9 männliche Studierende und 9 weibliche Studierende
Abiturnote	M = 2.2; SD = 0.54; Max. = 3.3; Min. = 1.5
Fächer, welche die Studierenden in der Oberstufe durchgehend belegt haben	Ein Studierender machte hierzu keine Angaben. 22.2 Prozent: Biologie 5.6 Prozent: Biologie und Physik 11.1 Prozent: Chemie 38.9 Prozent: Chemie und Biologie 16.7 Prozent: Chemie und Physik

Zweit- bzw. Drittfach, welches die Studierenden zum Fach Chemie gewählt haben	Acht Seminarteilnehmer studierten zusätzlich das Fach Biologie. Zehn Seminarteilnehmer studierten Mathematik. Aus letzter Gruppe hatte ein Studierender das Drittfach Deutsch als Fremdsprache und ein Studierender das Drittfach Physik.
Qualität der Hochschulausbildung wurde mit Schulnoten wie folgt bewertet	fachliche Ausbildung: $M = 2,44$; $SD = 0.71$; $Max. = 4$; $Min. = 1$ fachdidaktische Ausbildung: $M = 1,78$; $SD = 0.81$; $Max. = 4$; $Min. = 1$ psych./päd. Ausbildung: $M = 2,67$; $SD = 0.91$; $Max. = 4$; $Min. = 1$

Die restlichen Teile des Fragebogens bestanden aus dem Oberthema Zielorientierungen, dem Oberthema fachspezifische Überzeugungen mit den Unterpunkten Chemie als Wissenschaft und Lehren und Lernen von Chemie in der Schule sowie dem Oberthema Selbstwirksamkeitserwartung, welches unterteilt wurde in Selbstwirksamkeitserwartung bezüglich der Planung von Experimenten, bezüglich der Durchführung von Experimenten, bezüglich der didaktischen Reduktion und bezüglich der didaktischen Rekonstruktion. Hierbei mussten die Studierenden die Items auf einer fünfstufigen Likert-Skala von „sehr zutreffend“ (entspricht dem Wert 1) bis „sehr unzutreffend“ (entspricht dem Wert 5) bewerten; ein niedriger Wert bedeutet eine positive Ausprägung. Zunächst wurden negativ formulierte Items recodiert und im Anschluss die Reliabilitäten der Pilotstudie berechnet.

Studierendenfragebogen – Zielorientierungen (Tabelle 57)

In dem Bereich selbstbezogene Ziele der Teilnehmer in ihrem Studium wurde eine mäßige Reliabilität von .67 bei 16 Items berechnet. Als Kontrollvariable erachtet Rost (2013) dies als akzeptabel. Die Ergebnisse werden somit nur als Tendenzen ausgelegt.

Tabelle 57: Deskriptive Statistik zu Zielorientierungen über drei Messzeitpunkte – Studierendenfragebogen – Pilotstudie gesamt

Pilotstudie gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	18	1.87	0.31	2.69	1.38
Post-Test	18	1.90	0.35	2.56	1.06
Follow-Up-Test	18	1.95	0.39	2.88	1.31

Aufteilung in Gruppen (Tabelle 58, Tabelle 59, Tabelle 60):

Tabelle 58: Deskriptive Statistik zu Zielorientierungen über drei Messzeitpunkte – Studierendenfragebogen – Feedbackgruppe

Feedbackgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	1.83	0.36	2.19	1.38
Post-Test	6	1.90	0.29	2.31	1.50
Follow-Up-Test	6	1.81	0.39	2.38	1.31

Tabelle 59: Deskriptive Statistik zu Zielorientierungen über drei Messzeitpunkte – Studierendenfragebogen – Selbstreflexionsgruppe

Selbstreflexionsgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.00	0.40	2.69	1.56
Post-Test	6	1.79	0.48	2.56	1.06
Follow-Up-Test	6	2.00	0.52	2.88	1.38

Tabelle 60: Deskriptive Statistik zu Zielorientierungen über drei Messzeitpunkte – Studierendenfragebogen – Kontrollgruppe

Kontrollgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	1.77	0.06	1.88	1.69
Post-Test	6	2.02	0.26	2.44	1.75
Follow-Up-Test	6	2.03	0.24	2.38	1.75

Während sich die Werte der Feedbackgruppe und der Selbstreflexionsgruppe wenig ändern, verschlechtern sich die Werte der Kontrollgruppe tendenziell.

Studierendenfragebogen – Fachspezifische Überzeugungen

Der dritte Teil des Fragebogens beinhaltete Subskalen zu den fachspezifischen Überzeugungen. Hier konnten die 10 Items im Bereich Chemie als Wissenschaft und die 8 Items für den Bereich Lernen von Chemie in der Schule nur unzureichende Reliabilitäten erreichen (siehe Anhang). Für die zweite Subskala, Lehren von Chemie (Tabelle 61), wurde eine Reliabilität von .78 (21 Items) berechnet.

Studierendenfragebogen – Fachspezifische Überzeugungen – Lehren von Chemie in der Schule

Tabelle 61: Deskriptive Statistik zur Subskala Lehren von Chemie in der Schule über drei Messzeitpunkte – Fachspezifische Überzeugungen – Studierendenfragebogen – Pilotstudie gesamt

Pilotstudie gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	18	2.04	0.29	2.62	1.57
Post-Test	18	2.01	0.26	2.62	1.33
Follow-Up-Test	18	1.99	0.25	2.38	1.52

Aufteilung in Gruppen (Tabelle 62, Tabelle 63, Tabelle 64):

Tabelle 62: Deskriptive Statistik – Subskala Lehren von Chemie in der Schule über drei Messzeitpunkte – Fachspezifische Überzeugungen – Studierendenfragebogen – Feedbackgruppe

Feedbackgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	1.83	0.23	2.24	1.57
Post-Test	6	1.86	0.28	2.14	1.33
Follow-Up-Test	6	1.83	0.20	2.05	1.52

Tabelle 63: Deskriptive Statistik – Subskala Lehren von Chemie in der Schule über drei Messzeitpunkte – Fachspezifische Überzeugungen – Studierendenfragebogen – Selbstreflexionsgruppe

Selbstreflexionsgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.14	0.22	2.43	1.81
Post-Test	6	2.09	0.13	2.19	1.86
Follow-Up-Test	6	2.02	0.25	2.24	1.52

Tabelle 64: Deskriptive Statistik – Subskala Lehren von Chemie in der Schule über drei Messzeitpunkte – Fachspezifische Überzeugungen – Studierendenfragebogen – Kontrollgruppe

Kontrollgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.14	0.32	2.62	1.81
Post-Test	6	2.10	0.31	2.62	1.71
Follow-Up-Test	6	2.13	0.21	2.38	1.81

Studierendenfragebogen – Fachspezifische Überzeugungen gesamt

Für alle Items im Bereich Fachspezifische Überzeugungen gesamt konnte eine Reliabilität von .82 (39 Items) berechnet werden (Tabelle 65).

Tabelle 65: Deskriptive Statistik – Fachspezifische Überzeugungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Pilotstudie gesamt

Fachsp. Überzeugungen gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	18	2.14	0.28	2.77	1.74
Post-Test	18	2.09	0.29	2.82	1.49
Follow-Up-Test	18	2.12	0.27	2.56	1.64

Aufteilung in Gruppen (Tabelle 66, Tabelle 67, Tabelle 68):

Tabelle 66: Deskriptive Statistik – Fachspezifische Überzeugungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Feedbackgruppe

Feedbackgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	1.90	0.15	2.10	1.74
Post-Test	6	1.85	0.20	2.05	1.49
Follow-Up-Test	6	1.89	0.25	2.26	1.64

Tabelle 67: Deskriptive Statistik – Fachspezifische Überzeugungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Selbstreflexionsgruppe

Selbstreflexionsgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.23	0.24	2.46	1.82
Post-Test	6	2.18	0.20	2.44	1.95
Follow-Up-Test	6	2.20	0.19	2.36	1.82

Tabelle 68: Deskriptive Statistik – Fachspezifische Überzeugungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Kontrollgruppe

Kontrollgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.30	0.29	2.77	1.90
Post-Test	6	2.23	0.32	2.82	1.85
Follow-Up-Test	6	2.27	0.21	2.56	2.08

Studierendenfragebogen – Selbstwirksamkeitserwartungen

Der letzte Teil des Studierendenfragebogens war unterteilt in vier Subskalen zum Oberthema Selbstwirksamkeitserwartungen. Folgende Reliabilitäten wurden berechnet (Tabelle 69):

Tabelle 69: Reliabilitäten – Selbstwirksamkeitserwartung

Selbstwirksamkeitserwartung bezüglich ...	<i>α</i>	Anzahl der Items
...der Planung von Experimenten	.67	5
...der Durchführung von Experimenten	.74	4
...der didaktischen Rekonstruktion unterrichtsrelevanter Inhalte	.77	8
...der didaktischen Rekonstruktion im Unterricht	.69	5
Selbstwirksamkeitserwartung – gesamt	.88	22

Bei den Selbstwirksamkeitserwartungen konnten im Post-Test teilweise nur 17 Studierende einbezogen werden, da ein Proband der Kontrollgruppe eine Seite des Tests nicht bearbeitet hatte. Diese Seite bezog sich auf die Subskalen Planung von Experimenten und Durchführung von Experimenten.

Studierendenfragebogen – Selbstwirksamkeitserwartungen – Subskala Planung von Experimenten (Tabelle 70)

Tabelle 70: Deskriptive Statistik – Subskala Planung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Pilotstudie gesamt

Pilotstudie – gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	18	2.68	0.42	3.40	1.80
Post-Test	17	2.40	0.49	3.40	1.60
Follow-Up-Test	18	2.42	0.49	3.20	1.40

Aufteilung in Gruppen:

Tabelle 71: Deskriptive Statistik – Subskala Planung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Feedbackgruppe

Feedbackgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.70	0.17	3.00	2.60
Post-Test	6	2.33	0.52	2.80	1.60
Follow-Up-Test	6	2.37	0.34	2.80	1.80

Tabelle 72: Deskriptive Statistik – Subskala Planung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Selbstreflexionsgruppe

Selbstreflexionsgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.43	0.56	3.20	1.80
Post-Test	6	2.37	0.50	3.20	1.80
Follow-Up-Test	6	2.20	0.64	3.00	1.40

Tabelle 73: Deskriptive Statistik – Subskala Planung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Kontrollgruppe

Kontrollgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.90	0.35	3.40	2.40
Post-Test	5	2.52	0.54	3.40	2.00
Follow-Up-Test	6	2.70	0.37	3.20	2.20

Die Subskala Planung von Experimenten verbessert sich in der Pilotstudie gesamt von Pre-Test auf Post-Test leicht. Auch in den einzelnen Gruppen ist eine tendenzielle Verbesserung erkennbar.

Studierendenfragebogen – Selbstwirksamkeitserwartungen – Subskala Durchführung von Experimenten (Tabelle 74)

Tabelle 74: Deskriptive Statistik – Subskala Durchführung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Pilotstudie gesamt

Pilotstudie – gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	18	2.61	0.43	3.50	1.75
Post-Test	17	2.31	0.58	3.00	1.00
Follow-Up-Test	18	2.33	0.59	3.25	1.25

Aufteilung in Gruppen (Tabelle 75, Tabelle 76, Tabelle 77):

Tabelle 75: Deskriptive Statistik – Subskala Durchführung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Feedbackgruppe

Feedbackgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.63	0.38	3.00	2.00
Post-Test	6	2.17	0.58	3.00	1.50
Follow-Up-Test	6	2.08	0.41	2.75	1.75

Tabelle 76: Deskriptive Statistik – Subskala Durchführung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Selbstreflexionsgruppe

Selbstreflexionsgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.42	0.44	3.00	1.75
Post-Test	6	2.13	0.68	3.00	1.00
Follow-Up-Test	6	2.17	0.75	3.00	1.25

Tabelle 77: Deskriptive Statistik – Subskala Durchführung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Kontrollgruppe

Kontrollgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.79	0.46	3.50	2.25
Post-Test	5	2.70	0.27	3.00	2.25
Follow-Up-Test	6	2.75	0.39	3.25	2.25

Auffällig ist, dass im Post-Test sowohl in der Feedbackgruppe als auch in der Selbstreflexionsgruppe ein tendenziell größerer Zugewinn an Selbstwirksamkeit in der Subskala Durchführung von Experimenten deutlich wird als in der Kontrollgruppe, welche als einzige Gruppe keine Unterrichtsstunden halten durfte.

Studierendenfragebogen – Selbstwirksamkeitserwartungen – Subskala Didaktische Rekonstruktion unterrichtsrelevanter Inhalte (Tabelle 78)

Tabelle 78: Deskriptive Statistik – Subskala Didaktische Rekonstruktion unterrichtsrelevanter Inhalte über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Pilotstudie gesamt

Pilotstudie – gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	18	2.50	0.43	3.25	1.75
Post-Test	18	2.30	0.38	3.00	1.63
Follow-Up-Test	18	2.28	0.48	3.00	1.38

Aufteilung in Gruppen (Tabelle 79, Tabelle 80, Tabelle 81):

Tabelle 79: Deskriptive Statistik – Subskala Didaktische Rekonstruktion unterrichtsrelevanter Inhalte über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Feedbackgruppe

Feedbackgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.33	0.39	2.88	1.75
Post-Test	6	2.15	0.31	2.63	1.75
Follow-Up-Test	6	2.00	0.35	2.38	1.38

Tabelle 80: Deskriptive Statistik – Subskala Didaktische Rekonstruktion unterrichtsrelevanter Inhalte über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Selbstreflexionsgruppe

Selbstreflexionsgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.42	0.44	3.00	1.75
Post-Test	6	2.25	0.43	2.88	1.63
Follow-Up-Test	6	2.19	0.49	2.75	1.50

Tabelle 81: Deskriptive Statistik – Subskala Didaktische Rekonstruktion unterrichtsrelevanter Inhalte über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Kontrollgruppe

Kontrollgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.75	0.40	3.25	2.25
Post-Test	6	2.50	0.34	3.00	2.13
Follow-Up-Test	6	2.67	0.35	3.00	2.25

Die Feedbackgruppe erzielte, wie die Selbstreflexionsgruppe, sowohl im Post-Test als auch im Follow-Up Test leicht bessere Punktwerte zur Subskala Didaktische Rekonstruktion unterrichtsrelevanter Inhalte. Lediglich die Kontrollgruppe weist zum Follow-Up-Test geringfügig weniger Punkte auf.

Studierendenfragebogen – Selbstwirksamkeitserwartungen – Subskala Didaktische Rekonstruktion im Unterricht

Tabelle 82: Deskriptive Statistik – Subskala Didaktische Rekonstruktion im Unterricht über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Pilotstudie gesamt

Pilotstudie - gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	18	2.36	0.64	3.60	1.40
Post-Test	18	2.31	0.49	3.20	1.40
Follow-Up-Test	18	2.14	0.47	3.00	1.40

Aufteilung in Gruppen:

Tabelle 83: Deskriptive Statistik – Subskala Didaktische Rekonstruktion im Unterricht über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Feedbackgruppe

Feedbackgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.17	0.37	2.60	1.60
Post-Test	6	2.30	0.62	3.20	1.60
Follow-Up-Test	6	1.90	0.33	2.20	1.40

Tabelle 84: Deskriptive Statistik – Subskala Didaktische Rekonstruktion im Unterricht über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Selbstreflexionsgruppe

Selbstreflexionsgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.37	0.66	3.20	1.80
Post-Test	6	2.10	0.47	2.80	1.40
Follow-Up-Test	6	1.93	0.39	2.40	1.40

Tabelle 85: Deskriptive Statistik – Subskala Didaktische Rekonstruktion im Unterricht über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Kontrollgruppe

Kontrollgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.53	0.86	3.60	1.40
Post-Test	6	2.53	0.33	3.00	2.00
Follow-Up-Test	6	2.60	0.36	3.00	2.00

Es lässt sich eine tendenzielle Verbesserung der Subskala Didaktische Rekonstruktion im Unterricht erkennen. Hierbei scheint diese von Post-Test auf Follow-Up-Test größer zu sein als von Pre-Test auf Post-Test. Lediglich die Kontrollgruppe verschlechtert sich leicht.

Studierendenfragebogen – Selbstwirksamkeitserwartungen gesamt (Tabelle 86)*Tabelle 86: Deskriptive Statistik – Selbstwirksamkeitserwartungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Pilotstudie gesamt*

Pilotstudie – gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	18	2.53	0.40	3.32	1.91
Post-Test	18	2.32	0.36	2.86	1.50
Follow-Up-Test	18	2.29	0.47	3.00	1.45

Im Allgemeinen haben sich die Selbstwirksamkeitserwartungen der Probanden der Pilotstudie tendenziell verbessert.

Es folgt die Aufteilung in Gruppen (Tabelle 87, Tabelle 88, Tabelle 89):

Tabelle 87: Deskriptive Statistik – Selbstwirksamkeitserwartungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Feedbackgruppe

Feedbackgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.43	0.24	2.68	2.14
Post-Test	6	2.23	0.28	2.64	1.82
Follow-Up-Test	6	2.08	0.32	2.45	1.55

Tabelle 88: Deskriptive Statistik – Selbstwirksamkeitserwartungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Selbstreflexionsgruppe

Selbstreflexionsgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.41	0.47	3.09	1.91
Post-Test	6	2.22	0.44	2.68	1.50
Follow-Up-Test	6	2.13	0.54	2.68	1.45

Tabelle 89: Deskriptive Statistik – Selbstwirksamkeitserwartungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Kontrollgruppe

Kontrollgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	6	2.74	0.42	3.32	2.36
Post-Test	6	2.51	0.31	2.86	2.08
Follow-Up-Test	6	2.67	0.31	3.00	2.23

Während sowohl in der Feedbackgruppe als auch in der Selbstreflexionsgruppe die Selbstwirksamkeitserwartung stetig steigt, ist in der Kontrollgruppe keine stetige Steigerung zu beobachten.

6.1.5. Kodiermanual

Um die Reliabilität einer qualitativen Inhaltsanalyse zu interpretieren, müssen die Punkte Stabilität (Intracoderreliabilität), Reproduzierbarkeit (Intercoderreliabilität) und Exaktheit, welche Stabilität und Reproduzierbarkeit voraussetzt, berücksichtigt werden.

Stabilität (Intracoderreliabilität)

Um die Intracoderreliabilität zu ermitteln, kodierte die Leiterin der Studie 10 Prozent der Messzeitpunkte erneut. Dies erfolgte in einem zeitlichen Abstand von ca. zwölf Monaten. Im Anschluss wurden die prozentuale Übereinstimmung, der Holsti-Koeffizient und Krippendorffs Alpha bestimmt.

Um die prozentuale Übereinstimmung und den Holsti-Koeffizienten zu ermitteln, wurden zunächst mithilfe des Statistikprogramms SPSS die Differenzen der einzelnen Kodierungen errechnet. Anschließend wurden mittels deskriptiver Häufigkeiten die Differenzen mit dem Wert 0 gezählt. Diese stehen für die übereinstimmenden Kodierungen. Bei der prozentualen Übereinstimmung wurde schließlich durch die Anzahl der Kodierungen dividiert. Hierbei zählen auch einseitige Kodierungen doppelt. Der Unterschied bei der Berechnung des Holsti-Koeffizienten ist, dass die Anzahl der Übereinstimmungen durch die tatsächliche Anzahl der Kodierungen der beiden Kodierer geteilt wird. Nichtkodierungen gehen also nicht in die Rechnung mit ein. Krippendorffs Alpha wurde mithilfe eines im Internet frei verfügbaren Makros von SPSS direkt ausgegeben (Tabelle 90):

Tabelle 90: Berechnungen zur Intracoderreliabilität

Berechnung Intracoderreliabilität	Wert
Prozentuale Übereinstimmung	70 %
Holsti-Koeffizient	$c_r = .71$
Krippendorffs Alpha	$\alpha = .68$

Diese Werte werden als gute Übereinstimmungen interpretiert.

Reproduzierbarkeit (Intercoderreliabilität)

Um die Reproduzierbarkeit des Kodiermanuals überprüfen zu können, wurden insgesamt drei Kodierer in das Projekt miteinbezogen. Alle drei Kodierer waren auch Teil dieser Studie. Kodierer 1 war zusätzlich Proband in der Pilotstudie und erhielt somit Input über Reflexionskompetenz und darüber, was gute Reflexionen auszeichnen. Ebenso erstellte Kodierer 1 eine Zulassungsarbeit zum Thema Anpassung eines Kodiermanuals zur Messung der Entwicklung der Reflexionskompetenz. Kodierer 2 war als Feedbackgeber der gesamten Studie für die erste Interventionsgruppe zuständig. Auch ihm boten sich aufgrund seiner Zulassungsarbeit zum Thema Feedback zusätzliche Möglichkeiten, sich sehr gut in das Thema und die Studie einzuarbeiten. Kodierer 3 bediente während der Schülertage eine Kamera und hatte somit mehrere Vergleichsmöglichkeiten der Stunden und zudem Einsicht in die Entwicklung der Stunden über drei Messzeitpunkte. Ebenso war Kodierer 3 für viele Testungen der Schülerinnen und Schüler zuständig. Eine explizite Einarbeitung in das Thema Reflexion fand bei Kodierer 3 nicht statt.

1. Arbeitsschritt (Kodierer 1):

Das Grundgerüst erarbeitete die Leiterin der Studie. Um eine zufriedenstellende Reliabilität des Kodiermanuals zu erreichen, wurden in Zusammenarbeit mit Kodierer 1 mehrfach Kategorien umgeordnet, neu definiert oder zusammengefasst. So wurden die Kategorien der Reflexionsbreite induktiv aus den schriftlichen Reflexionen der Probanden der Pilotstudie abgeleitet. Deduktiv dienten die Merkmale guten Unterrichts von Helmke (2009) und Meyer (2003) sowie die Metastudie von Hattie (2008) als Anhaltspunkte zur

Kategorienbildung. Da den einzelnen Kategorien nicht immer die gleichen Inhalte zugeordnet wurden, erfolgte eine induktive Prüfung am Material. Die Reflexionstiefe wurde deduktiv aus diversen Quellen erarbeitet (Kapitel 3.4). Hierbei waren die Oberpunkte Beschreibung, Bewertung, Alternative und Konsequenz bereits von Anfang an im Manual vertreten. Alle Oberpunkte ließen sich zunächst verschiedenartig begründen. Die Unterteilung hierbei beinhaltete allgemeine Begründungen, ein Abwägen verschiedener Vor- und Nachteile, die alleinige Nennung von Vorteilen sowie die ledigliche Nennung von Nachteilen und Zielbegründungen. Die gewählten Kategorien wurden sowohl deduktiv als auch induktiv entwickelt, mehrfach diskutiert, überarbeitet und überprüft. Im Rahmen einer Zulassungsarbeit wurden gute Werte in der Reflexionsbreite erreicht. Die Reflexionstiefe war nicht zufriedenstellend (Riebesecker, 2018).

2. Arbeitsschritt (Kodierer 2):

Wie sich herausstellte, wurden die Definitionen der Reflexionstiefe nicht exakt genug gegeben, um den Kodierern ein übereinstimmendes Arbeiten zu ermöglichen. Deshalb wurden nicht trennbare Kategorien zusammengefügt oder Kodierregeln, Ankerbeispiele und Abgrenzungen ergänzt und optimiert. Die Unterteilung der vier Oberkategorien wurde jeweils auf Rechtfertigung, Abwägung und Zielformulierung begrenzt. Zusätzlich zur Reflexionstiefe wurden Perspektivenübernahmen gewertet. Zunächst erfolgte eine Kodierung der drei Perspektiven Schülerperspektive, Theorieperspektive und Fachperspektive neben der allgemein eingenommenen Lehrerperspektive. Diese wurden jedoch im Verlauf der Auswertungen induktiv um die Politisch-Soziale-Perspektive ergänzt. Zusätzlich wurden Literaturbezüge kodiert. Diese zweite kommunikative Validierung mit Kodierer 2 half dabei, die Gültigkeit der Kategorien der Reflexionstiefe besser einzuschätzen. In einer ersten Übungsphase wurden gemeinsam 3 Reflexionsbeispiele kodiert. Hierbei wurden erneut die Regeln des Manuals beproben und etwaige Unklarheiten überarbeitet. Im Anschluss wurden 10 Prozent der Pilotstudie separat von Kodierer 2 und der Studienleiterin kodiert. Folgende Werte wurden berechnet (Tabelle 91):

Tabelle 91: Berechnungen zur Interkoderreliabilität – Pilotstudie gesamt – Kodierer 2

Interkoderreliabilität Kodierer 2 Pilotstudie	Wert
Prozentuale Übereinstimmung	76 %
Holsti-Koeffizient	$c_r = .77$
Krippendorffs Alpha	$\alpha = .73$

Berechnung der Reliabilitäten der Subskalen des Kodiermanuals (Tabelle 92):

Tabelle 92: Berechnungen zur Interkoderreliabilität – Pilotstudie Subskalen – Kodierer 2

Interkoderreliabilität Kodierer 2 Pilotstudie	Reflexionsbreite	Reflexionstiefe
Prozentuale Übereinstimmung	81 %	75 %
Holsti-Koeffizient	$c_r = .81$	$c_r = .76$
Krippendorffs Alpha	$\alpha = .78$	$\alpha = .70$

Nach der Berechnung dieser bereits als sehr gut interpretierbaren Werte wurden die kodierten Messzeitpunkte der Pilotstudie erneut durchgesprochen. Im Manual wurden kleinere Kodierregeln ergänzt und zwei Definitionen umformuliert. Mit dem finalen Manual wurden anschließend 10 Prozent der Hauptstudie doppelt kodiert. Da sich der Holsti-Koeffizient und der Wert für Krippendorfs Alpha in der Pilotstudie lediglich gering unterschieden haben, wird im Folgenden auf den Holsti-Koeffizienten verzichtet, da dieser eine zufällige Übereinstimmung der Kodierer vernachlässigt (Tabelle 93).

Tabelle 93: Berechnungen zur Interoderreliabilität – Hauptstudie gesamt – Kodierer2

Interoderreliabilität Kodierer 2 Hauptstudie	Wert
Prozentuale Übereinstimmung	86 %
Krippendorfs Alpha	$\alpha = .88$

Berechnung der Reliabilitäten der Subskalen des Kodiermanuals (Tabelle 94):

Tabelle 94: Berechnungen zur Interoderreliabilität – Hauptstudie Subskalen – Kodierer 2

Interoderreliabilität Kodierer 2 Hauptstudie	Reflexionsbreite	Reflexionstiefe
Prozentuale Übereinstimmung	88 %	85 %
Krippendorfs Alpha	$\alpha = .86$	$\alpha = .88$

3. Arbeitsschritt (Kodierer 3):

Da sowohl Kodierer 1 als auch Kodierer 2 Teil der Studie wie auch bei der Erarbeitung des Manuals miteinbezogen waren und selbst Arbeiten zum Thema Reflexionskompetenz verfasst hatten, wurde mit Kodierer 3 eine Person gewählt, welche ebenfalls an der Studie mitarbeitete, die aber weder Teil der Entwicklung des Manuals noch in das Thema Reflexionskompetenz eingearbeitet war. Somit hatte Kodierer 3 weniger Übung beim Kodieren und Verstehen eines Manuals und musste sich zunächst wesentlich übungsintensiver einlesen und einarbeiten. Nach einer gemeinsamen Übungsphase wurden 10 Prozent der Pilotstudie kodiert (Tabelle 95, Tabelle 96):

Tabelle 95: Berechnungen zur Interoderreliabilität – Pilotstudie gesamt – Kodierer 3

Interoderreliabilität Kodierer 3 Pilotstudie	Wert
Prozentuale Übereinstimmung	63 %
Krippendorfs Alpha	$\alpha = .61$

Tabelle 96: Berechnungen zur Interoderreliabilität – Pilotstudie Subskalen – Kodierer 3

Interoderreliabilität Kodierer 3 Pilotstudie	Reflexionsbreite	Reflexionstiefe
Prozentuale Übereinstimmung	75 %	61 %
Krippendorfs Alpha	$\alpha = .71$	$\alpha = .57$

Die Werte sind als akzeptable bis gute Übereinstimmung interpretierbar. Da die Werte jedoch deutlich geringer ausfielen als bei Kodierer 2, wurden mit Kodierer 3 ebenfalls alle Messzeitpunkte der Pilotstudie genau durchgesprochen. Das Manual wurde nicht mehr verändert. Besonderen Wert wurde dabei auf die Ankerbeispiele und ihre Auslegung gelegt. Im Anschluss wurden 10 Prozent der Hauptstudie kodiert (Tabelle 97, Tabelle 98):

Tabelle 97: Berechnungen zur Intercoderreliabilität – Hauptstudie gesamt – Kodierer 3

Intercoderreliabilität Kodierer 3 Hauptstudie	Wert
Prozentuale Übereinstimmung	80 %
Krippendorffs Alpha	$\alpha = .88$

Tabelle 98: Berechnungen zur Intercoderreliabilität – Hauptstudie Subskalen – Kodierer 3

Intercoderreliabilität Kodierer 3 Hauptstudie	Reflexionsbreite	Reflexionstiefe
Prozentuale Übereinstimmung	88 %	79 %
Krippendorffs Alpha	$\alpha = .86$	$\alpha = .88$

Die intensive Besprechung der Pilotstudie führte auch bei Kodierer 3 zu sehr guten Übereinstimmungen. Kember et al. (1999) empfehlen, projektinterne Kodierer einzusetzen. Dies wurde in dieser Studie versucht, trotzdem waren die Voraussetzungen unterschiedliche, welche der Situation entsprechend berücksichtigt wurden. So fiel die Besprechung der Pilotstudienresultate bei Kodierer 3 deutlich länger aus als bei Kodierer 2. Ebenso wurde nur mit Kodierer 3 das Manual intensiver besprochen, da Kodierer 2 bei der Entwicklung beteiligt war.

Exaktheit

Die Exaktheit lässt sich nach Krippendorff (2018) in vier Bereiche unterteilen: Auswertungseinheit, Analytiker, einzelne Kategorien und Kategoriendifferenzierungen. Es wurde mit dem frei verfügbaren Programm RQDA auf systematische Fehler in den Auswertungseinheiten hin überprüft. Hierbei wurden die Erstkodierungen mit den Nachkodierungen ausgegeben und die deskriptiven Häufigkeiten der fehlerhaft ausgewählten Kategorien überprüft. Hierbei ergaben sich keine Auffälligkeiten. Ebenso wurden die Kategorien überprüft und bei auffälligen Unstimmigkeiten die Definitionen überarbeitet. Die Intercoderreliabilitäten wurden bereits ermittelt und uneindeutige Kategorien zu größeren und so exakter anwendbaren Kategorien zusammengefügt (Arbeitsschritt 2) (Mayring, 2010). Somit kann von einer ausreichenden Reliabilität ausgegangen werden. Die deskriptiven Analysen ergeben folgende Werte:

Summe der insgesamt erreichten Punkte über drei Messzeitpunkte der gesamten Pilotstudie (Abbildung 45, Tabelle 99):

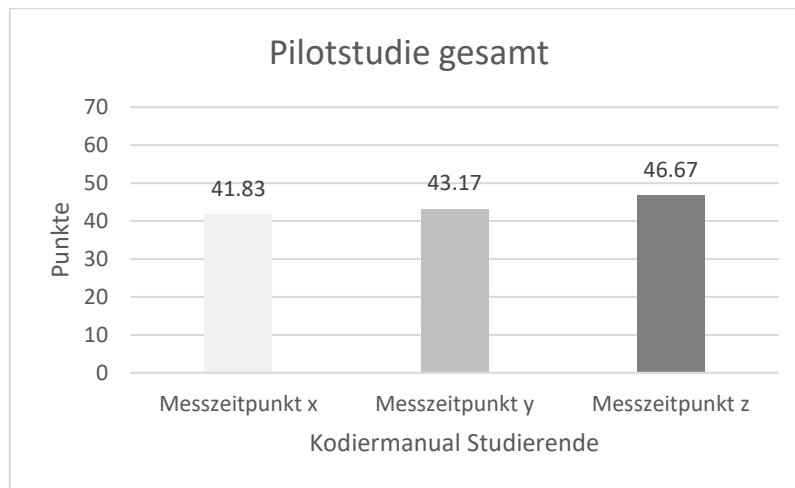


Abbildung 45: Gesamtpunktzahl Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z

Tabelle 99: Deskriptive Statistik – Gesamtpunktzahl Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z

	N	M	SD	Max.	Min.
Messzeitpunkt x	18	41.83	14.05	73.00	12.00
Messzeitpunkt y	18	43.17	17.65	91.00	11.00
Messzeitpunkt z	18	46.67	13.42	72.00	14.00

Summe der insgesamt erreichten Punkte aufgeteilt in Gruppen:

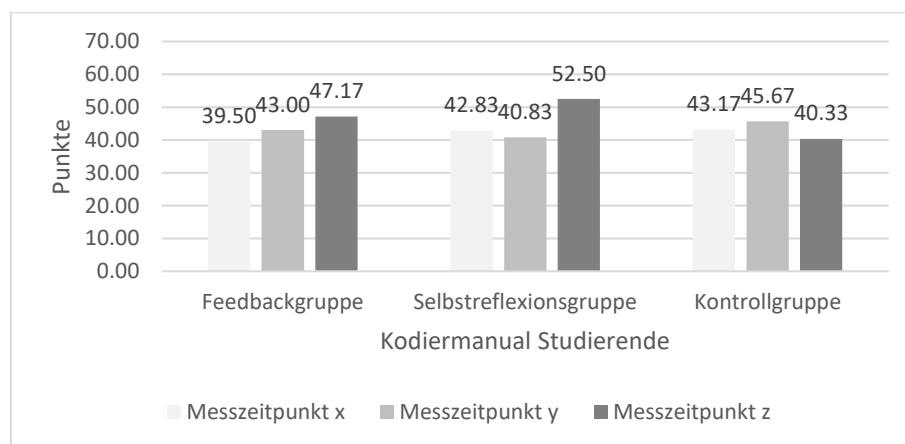


Abbildung 46: Gesamtpunktzahl Kodiermanual Pilotstudie – Punkte im Gruppenvergleich – Messzeitpunkte x,y,z

Tabelle 100: Deskriptive Statistik – Gesamtpunktzahl Kodiermanual Pilotstudie – Punkte Feedbackgruppe – Messzeitpunkte x,y,z

Feedbackgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	6	39.50	20.40	73.00	12.00
Messzeitpunkt y	6	43.00	18.79	64.00	11.00
Messzeitpunkt z	6	47.17	12.45	61.00	28.00

Tabelle 101: Deskriptive Statistik – Gesamtpunktzahl Kodiermanual Pilotstudie – Punkte Selbstreflexionsgruppe – Messzeitpunkte x,y,z

Selbstreflexionsgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	6	42.83	11.14	61.00	33.00
Messzeitpunkt y	6	40.83	13.42	58.00	29.00
Messzeitpunkt z	6	52.50	13.41	72.00	34.00

Tabelle 102: Deskriptive Statistik – Gesamtpunktzahl Kodiermanual Pilotstudie – Punkte Kontrollgruppe – Messzeitpunkte x,y,z

Kontrollgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	6	43.17	11.00	59.00	33.00
Messzeitpunkt y	6	45.67	22.62	91.00	29.00
Messzeitpunkt z	6	40.33	13.71	50.00	14.00

Insgesamt betrachtet ist in den Interventionsgruppen vom ersten zum dritten Messzeitpunkt eine Steigerung zu beobachten. Die Kontrollgruppe fällt jedoch ab.

Im Folgenden werden die Subskalen Reflexionsbreite und Reflexionstiefe zu den drei Messzeitpunkten x (Reflexion nach erster Stundenplanung und ggf. -durchführung), y (Reflexion nach zweiter Stundenplanung und ggf. -durchführung) und z (Reflexion nach dritter Stundenplanung und ggf. -durchführung).

Reflexionsbreite (Abbildung 47, Tabelle 103)

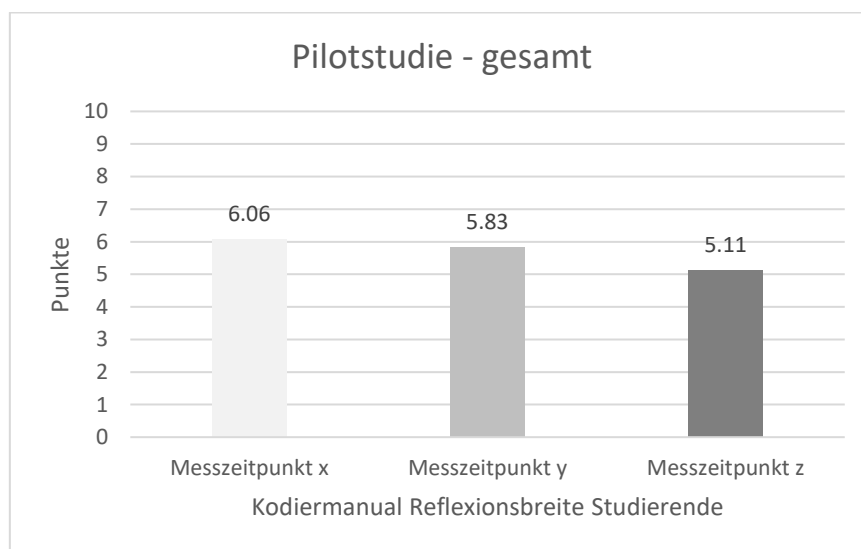


Abbildung 47: Reflexionsbreite Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z

Tabelle 103: Deskriptive Statistik – Reflexionsbreite Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z

Reflexionsbreite Pilotstudie gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	18	6.06	1.66	9.00	3.00
Messzeitpunkt y	18	5.83	1.47	8.00	2.00
Messzeitpunkt z	18	5.11	1.68	7.00	1.00

Aufteilung in Gruppen (Abbildung 48, Tabelle 104, Tabelle 105, Tabelle 106):

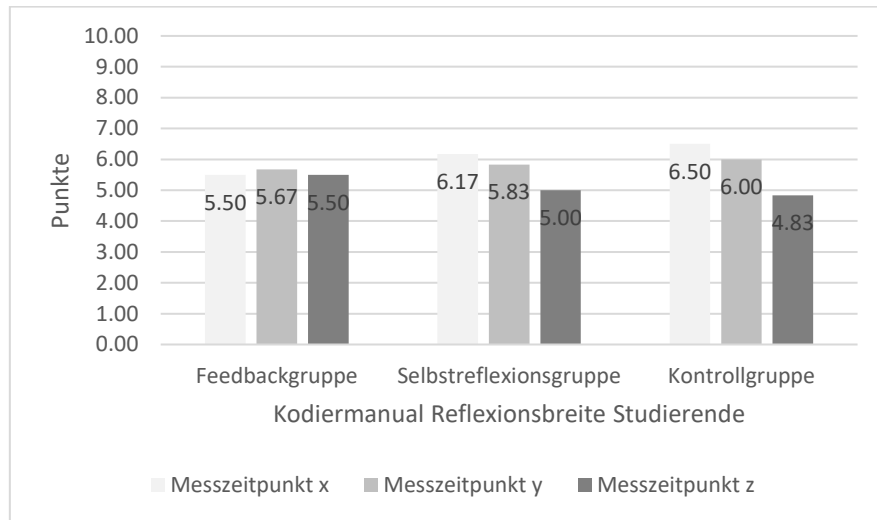


Abbildung 48: Reflexionsbreite Kodiermanual – Punkte im Gruppenvergleich – Messzeitpunkte x,y,z

Tabelle 104: Deskriptive Statistik – Reflexionsbreite Kodiermanual – Punkte Feedbackgruppe – Messzeitpunkte x,y,z

Reflexionsbreite Feedbackgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	6	5.50	1.38	7.00	3.00
Messzeitpunkt y	6	5.67	2.07	8.00	2.00
Messzeitpunkt z	6	5.50	1.38	7.00	3.00

Tabelle 105: Deskriptive Statistik – Reflexionsbreite Kodiermanual – Punkte Selbstreflexionsgruppe – Messzeitpunkte x,y,z

Reflexionsbreite Selbstreflexionsgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	6	6.17	1.33	8.00	4.00
Messzeitpunkt y	6	5.83	0.98	7.00	5.00
Messzeitpunkt z	6	5.00	1.41	7.00	3.00

Tabelle 106: Deskriptive Statistik – Reflexionsbreite Kodiermanual – Punkte Kontrollgruppe – Messzeitpunkte x,y,z

Reflexionsbreite Kontrollgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	6	6.50	2.26	9.00	3.00
Messzeitpunkt y	6	6.00	1.41	8.00	4.00
Messzeitpunkt z	6	4.83	2.32	7.00	1.00

Während die Reflexionsbreite in der Selbstreflexionsgruppe und Kontrollgruppe über die drei Messzeitpunkte stetig geringer wird, kann die Feedbackgruppe ihr Ergebnis zunächst verbessern, erreicht aber im dritten Messzeitpunkt wieder Ausgangsniveau.

Im Folgenden werden die Ergebnisse der Reflexionstiefe (Abbildung 49, Tabelle 107) der gesamten Pilotstudie dargestellt:

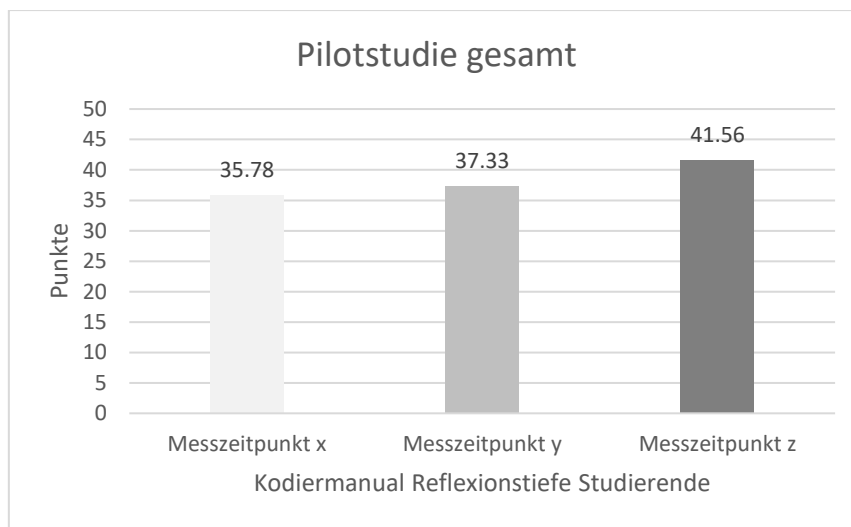


Abbildung 49: Reflexionstiefe Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z

Tabelle 107: Deskriptive Statistik – Reflexionstiefe Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z

Reflexionstiefe Pilotstudie gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	18	35.78	13.54	66.00	9.00
Messzeitpunkt y	18	37.33	16.64	83.00	9.00
Messzeitpunkt z	18	41.56	12.49	67.00	13.00

Es folgen die Ergebnisse der Reflexionstiefe der Pilotstudie aufgeteilt in Gruppen (Abbildung 50, Tabelle 108, Tabelle 109, Tabelle 110):

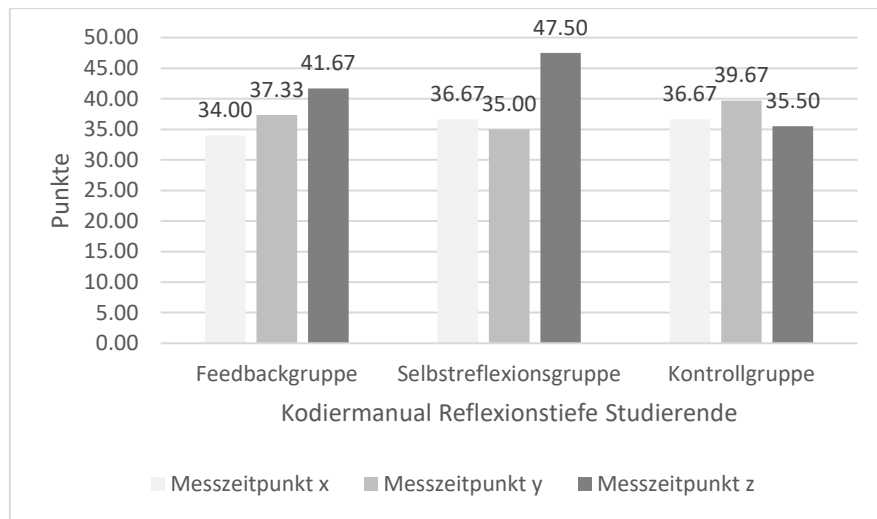


Abbildung 50: Reflexionstiefe Kodiermanual – Punkte im Gruppenvergleich – Messzeitpunkte x,y,z

Tabelle 108: Deskriptive Statistik – Reflexionstiefe Kodiermanual – Punkte Feedbackgruppe – Messzeitpunkte x,y,z

Reflexionstiefe Feedbackgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	6	34.00	19.14	66.00	9.00
Messzeitpunkt y	6	37.33	17.13	58.00	9.00
Messzeitpunkt z	6	41.67	11.57	55.00	25.00

Tabelle 109: Deskriptive Statistik – Reflexionstiefe Kodiermanual – Punkte Selbstreflexionsgruppe – Messzeitpunkte x,y,z

Reflexionstiefe Selbstreflexionsgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	6	36.67	11.78	55.00	27.00
Messzeitpunkt y	6	35.00	12.65	51.00	23.00
Messzeitpunkt z	6	47.50	12.88	67.00	31.00

Tabelle 110: Deskriptive Statistik – Reflexionstiefe Kodiermanual – Punkte Kontrollgruppe – Messzeitpunkte x,y,z

Reflexionstiefe Kontrollgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	6	36.67	10.61	51.00	24.00
Messzeitpunkt y	6	39.67	21.79	83.00	22.00
Messzeitpunkt z	6	35.50	12.01	46.00	13.00

Die Reflexionstiefe scheint gesamt betrachtet zu steigen. Bei der Unterteilung in die Gruppen wird jedoch deutlich, dass lediglich die Feedbackgruppe stetig an Punkten ge-

winnt. Die Selbstreflexionsgruppe reflektiert erst zum dritten Messzeitpunkt deutlich besser (tiefer). Die Kontrollgruppe verliert den Zugewinn vom ersten zum zweiten Messzeitpunkt im dritten Messzeitpunkt wieder.

6.2. Testinstrumente und Fragebogen – Schülerinnen und Schüler

6.2.1. Wissenstest – Schülerinnen und Schüler

Die Reliabilitätsanalyse des Wissenstests ergab zunächst nach der Pilotstudie ein Cronbachs Alpha von .56 bei 30 Items. Auch um die Schülerinnen und Schülern zu entlasten, wurden pro Unterthema ein bis zwei Items mit schlechten Trennschärfen gelöscht. Ebenso wurde in der Hauptstudie die Stunde zur Dichte aufgrund niedriger Studierendenzahlen nicht mehr unterrichtet, was zur Folge hatte, dass auch diese fünf Fragen aus dem Test genommen wurden. In der Hauptstudie 1 waren ferner in der Feedbackgruppe lediglich 4 Studierende, sodass nur vier Stunden gehalten und getestet wurden. Der finale Test bestand, mit Ausnahme der von den vier Studierenden der Feedbackgruppe in der Hauptstudie 1 unterrichteten Schülerinnen und Schülern, aus 17 Items. Hierbei konnte ein Cronbachs Alpha aus Pilot- und Hauptstudie von lediglich .50 nicht überzeugen. Da der Validität und damit der Breite des abzutestenden Wissens bei der Testkonzeption Priorität eingeräumt wurde, erscheint die relativ geringe Realibilität plausibel und vertretbar.

Es folgen die Ergebnisse der Pilotstudie des Wissenstests der Schülerinnen und Schüler (Abbildung 51, Tabelle 111):

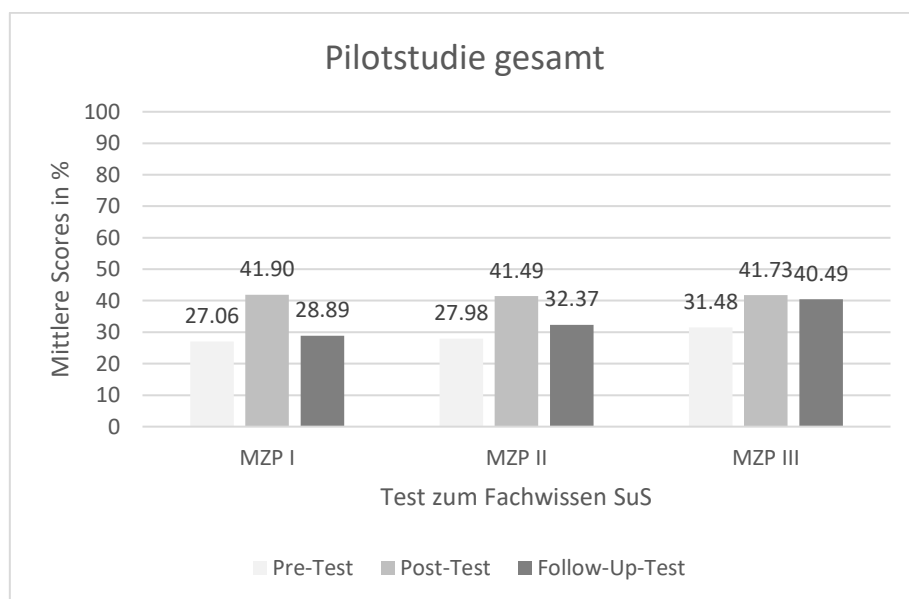


Abbildung 51: Mittlere Scores des Tests zum Fachwissen über drei Messzeitpunkte – SuS Pilotstudie gesamt

Tabelle 111: Deskriptive Statistik – Mittelwertvergleiche der Wissenstests über drei Messzeitpunkte – SuS Pilotstudie gesamt

Messzeitpunkt	Pilotstudie gesamt	N	M	SD	Max.	Min.
MZIP I	Pre-Test	42	27.06	8.72	43.33	3.33
	Post-Test	42	41.90	13.36	63.33	13.33

	Follow-Up-Test	42	28.89	13.29	70.00	10.00
MZP II	Pre-Test	38	27.98	10.78	53.33	3.33
	Post-Test	38	41.49	10.48	60.00	16.67
	Follow-Up-Test	38	32.37	10.42	56.67	13.33
MZP III	Pre-Test	27	31.48	9.62	46.67	6.67
	Post-Test	27	41.73	13.15	66.67	23.33
	Follow-Up-Test	27	40.49	10.32	70.00	23.33

Im Folgenden die Aufteilung der Ergebnisse des Wissenstests der Schülerinnen und Schüler, unterteilt in die Interventionsgruppen Feedback und Selbstreflexion (Abbildung 52, Abbildung 53, Tabelle 112, Tabelle 113):

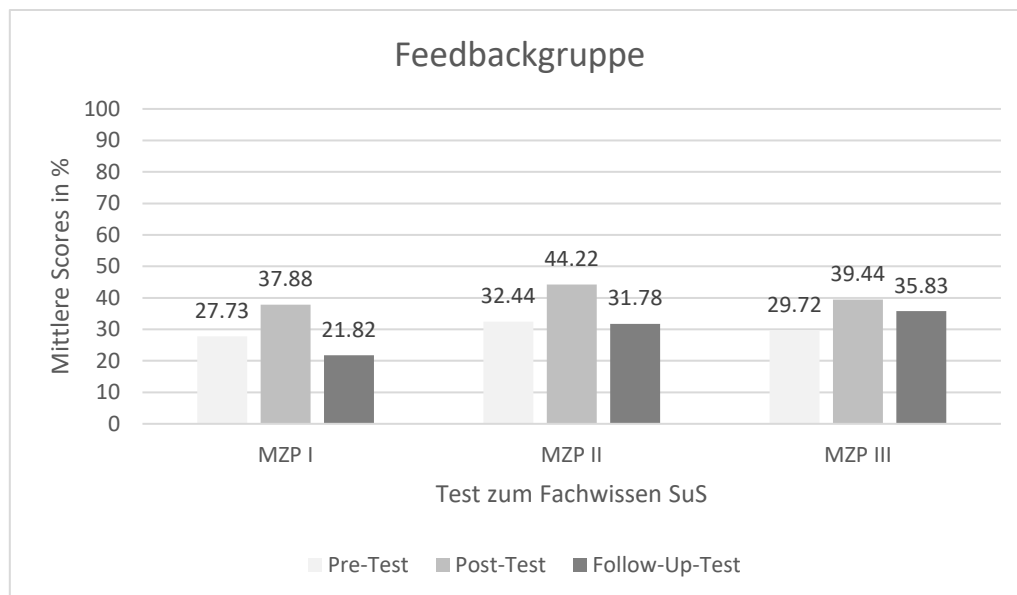


Abbildung 52: Mittlere Scores der Wissenstests über drei Messzeitpunkte – SuS Feedbackgruppe

Tabelle 112: Deskriptive Statistik – Mittelwertvergleiche der Wissenstests über drei Messzeitpunkte – SuS Feedbackgruppe

Messzeitpunkt	Feedbackgruppe	N	M	SD	Max.	Min.
MZP I	Pre-Test	22	27.73	8.93	40.00	3.33
	Post-Test	22	37.88	14.27	63.33	13.33
	Follow-Up-Test	22	21.82	7.54	36.67	10.00
MZP II	Pre-Test	15	32.44	9.13	53.33	16.67
	Post-Test	15	44.22	10.50	60.00	16.67

	Follow-Up-Test	15	31.78	9.25	50.00	16.67
MZP III	Pre-Test	12	29.72	10.29	43.33	6.67
	Post-Test	12	39.44	11.79	56.67	23.33
	Follow-Up-Test	12	35.83	7.80	50.00	23.33

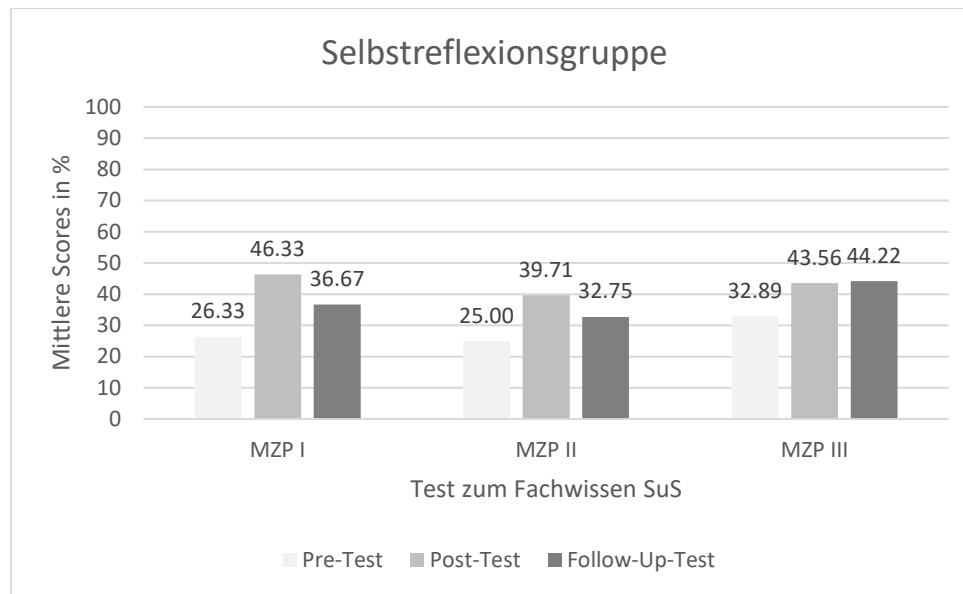


Abbildung 53: Mittlere Scores der Wissenstests über drei Messzeitpunkte – SuS Selbstreflexionsgruppe

Tabelle 113: Deskriptive Statistik – Mittelwertvergleiche der Wissenstests über drei Messzeitpunkte – SuS Selbstreflexionsgruppe

Messzeitpunkt	Selbstreflexionsgruppe	N	M	SD	Max.	Min.
MZP I	Pre-Test	20	26.33	8.65	43.33	10.00
	Post-Test	20	46.33	10.97	63.33	23.33
	Follow-Up-Test	20	36.67	14.02	70.00	13.33
MZP II	Pre-Test	23	25.00	10.96	50.00	3.33
	Post-Test	23	39.71	10.29	60.00	23.33
	Follow-Up-Test	23	32.75	11.31	56.67	13.33
MZP III	Pre-Test	15	32.89	9.16	46.67	16.67
	Post-Test	15	43.56	14.28	66.67	23.33
	Follow-Up-Test	15	44.22	10.80	70.00	26.67

6.2.2. Kognitiver Fähigkeitstest – Schülerinnen und Schüler

An den jeweils sechs Messzeitpunkten der Pilotstudie und der drei Hauptstudien wurden verschiedene Klassen unterrichtet. Um die Vergleichbarkeit der Klassen hinsichtlich ihrer kognitiven Fähigkeiten zu prüfen, kam der N2-Teil des kognitiven Fähigkeitentests von Heller und Perleth (2000) zum Einsatz. Dieser erreichte für die 25 überprüften Items in Pilot- und Hauptstudie ein Cronbachs Alpha von .86. Da die Stichprobe nach dem Kolmogorov-Smirnov Test nicht normalverteilt ist (siehe Anhang), wird zusätzlich zum T-Test für unabhängige Stichproben der Mann-Whitney-U-Test für nichtparametrische Stichproben angegeben. Damit konnte gezeigt werden, dass sich die Schülerinnen und Schüler der Feedbackgruppe im Hinblick auf die Kontrollvariable nicht signifikant von den Schülerinnen und Schülern der Selbstreflexionsgruppe unterscheiden. Die Effektstärken wurden nach Cohen (1988, 1992) berechnet. Die genauen Einteilungen werden in Tabelle 40 und in Tabelle 41 vorgestellt.

T-Test bei unabhängigen Stichproben: $t(480) = -1.07$; $p = .284$; n. s.; $d = 0.098$

Mann-Whitney-U-Test: $U(n_1=240, n_2 = 242) = -1.19$; $p = .234$; n. s.; $r = 0.05$

6.2.3. Fragebogen zum Fachinteresse – Schülerinnen und Schüler

Der Fragebogen zum Fachinteresse war unterteilt in die Subskalen Freude und Interesse an Chemie ($\alpha = .88$; 5 Items) und Chemiebezogenes Fähigkeitsselbstkonzept ($\alpha = .84$; 6 Items) und konnte sowohl in den Subskalen als auch gesamt ($\alpha = .90$; 11 Items) durch sehr gute Reliabilitäten überzeugen. Die Schülerinnen und Schüler entschieden hierbei auf einer sechsstufigen Likert-Skala von „stimme voll zu“ (entspricht dem Wert 1) bis „stimme gar nicht zu“ (entspricht dem Wert 6) (Tabelle 114).

Tabelle 114: Deskriptive Statistik zum Fachinteresse – Pilotstudie – Schülerinnen und Schüler

Pilotstudie	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Subskala: Freude und Interesse	107	3.03	1.05	5.60	1.00
Subskala: Fähigkeitsselbstkonzept	107	2.92	0.91	6.00	1.33
Fachinteresse gesamt	107	2.97	0.91	5.82	1.27

Die Werte können als durchschnittliches Interesse der Schülerinnen und Schüler für Chemie interpretiert werden.

Auch hier ließen sich keine Unterschiede bei den Schülerinnen und Schülern, bezogen auf die beiden Interventionsgruppen, finden:

$t(87,33) = 1.25$; $p = .213$; n. s.; $d = -0.246$

6.3. Konsequenzen für die Hauptstudie

Konsequenzen für den zeitlichen Umfang

Ursprünglich geplant waren pro Interventionsgruppe drei Schülertage zu je sechs Unterrichtsstunden, also sechs Probanden. Dies konnte aufgrund niedriger Studienzahlen während der Hauptstudie nicht realisiert werden. Zusätzlich waren die Schülerinnen und Schüler durch die sechs Stunden und den langen Aufenthalt an der Universität von 8:30 Uhr bis 15:00 Uhr stark gefordert. Um beiden Faktoren Rechnung zu tragen, wurden die Schülertage der Hauptstudie auf fünf Stunden reduziert, sodass die erste Stunde mit dem Thema Dichte entfiel. Diese Entscheidung fiel aufgrund der sehr physikalischen Umsetzung des Themas durch die Studierenden. Infolge verkürzte sich der Schülertag zeitlich um ca. eine Stunde (ca. 9:00 Uhr bis 14:45 Uhr).

Konsequenzen für die Tests

In der Pilotstudie wurden die Tests von der Studienleiterin und einer Mitarbeiterin sowie zusätzlich von drei mit der Studie vertrauten Lehrern umgesetzt. Diese wurden instruiert und konnten die Tests den Vorgaben entsprechend umsetzen. Um jedoch mögliche Fehlerquellen zu reduzieren, wurden in der Hauptstudie die Testung der Schülerinnen und Schüler und auch der Studierenden ausschließlich durch die Mitarbeiterin und die Studienleiterin instruiert und durchgeführt. Ebenso wurden folgende Tests aufgrund negativer Trennschärfen mancher Items und folglich niedriger Reliabilitäten dem Studiendesign angepasst:

Schülerinnen und Schüler

Der Wissenstest wurde in der Hauptstudie um die Items des Themas Dichte gekürzt. Ebenso wurden ein bis zwei Items der anderen Themengebiete gelöscht, so dass sowohl eine passende Reliabilität als auch die Validität gewährleistet blieben. An den beiden anderen Tests gab es keinerlei Änderungen.

Studierende

Der Studierendenfragebogen und auch der Test zum fachdidaktischen Wissen blieben unverändert. Der Fachwissenstest wurde dem Studiendesign angepasst und somit um die Items 7 und 13 reduziert. Ebenso wurde der Reflexionskompetenztest um die Items 12, 15, 17 und 18 reduziert.

Insgesamt erwies sich der Aufbau des Seminars und der Schülertage als stimmig und wurde gegenüber der Hauptstudie nicht weiter verändert. Somit sind Pilotierung und Hauptstudie hinsichtlich der Testungen und Durchführung grundsätzlich vergleichbar.

7. Ergebnisse zur Hauptstudie

In diesem Kapitel zur Hauptstudie folgt eine Darstellung der Ergebnisse unter Bezugnahme auf die gestellten zwei Forschungsfragen. Dabei werden diese mithilfe der erarbeiteten Hypothesen (Kapitel 4.2) aufgeklärt. Zunächst wird auf die Frage eingegangen, wie es möglich ist, ein universitäres Seminar zu gestalten, um die Reflexionskompetenz maximal zu fördern (Kapitel 7.1). Im Anschluss werden Zusammenhänge der Förderung der Reflexionskompetenz und dem fachdidaktischen Wissen dargestellt, ebenso wie Zusammenhänge mit dem Fachwissen, der Selbstwirksamkeit und der Unterrichtsqualität (Kapitel 7.2). Details zu den statistischen Auswertungen finden sich in Kapitel 10.3. Bei allen Berechnungen werden die Probanden der Pilotstudie mit den Probanden der Hauptstudie zusammengefasst. Es wurde vorab geprüft, ob sich die Daten der Pilot- von denen der Hauptstudie unterscheiden. Dazu wurden mit den zentralen Variablen Reflexionskompetenz, Fachwissen, fachdidaktisches Wissen sowie Überzeugungen und Selbstwirksamkeitserwartungen Varianzanalysen gerechnet. Als abhängige Variable wurde jeweils der Posttest-Wert und als unabhängige Variable der Pretest-Wert eingesetzt. Als Faktor wurde der Studienzeitpunkt (Pilotstudie, Hauptstudie) angegeben. Es konnten keine signifikanten Unterschiede zwischen Pilot- und Hauptstudie gefunden werden (siehe Anhang Unterkapitel 10.3.11). Es wurden somit im Folgenden die Probanden aus Pilot- und Hauptstudie zusammengefasst. Zur Hypothesentestung werden die Daten der gesamten Stichprobe herangezogen.

7.1. Ergebnisse zur Gestaltung eines universitären Seminars zur Förderung der Reflexionskompetenz (Forschungsfrage 1)

Der Aufbau des universitären Seminars erfolgte auf Basis des Projekts von Anthofer (2017). Um die Effektivität verschiedener Methoden auf den Schwerpunkt Reflexionskompetenz genauer zu untersuchen, wurden in dieser Studie die Probanden in zwei Interventionsgruppen und eine Kontrollgruppe aufgeteilt. Zur Überprüfung der Reflexionskompetenz wurde ein Wissenstest zum Thema Reflexion konzipiert, welcher sowohl deklaratives als auch konditionales Wissen abfragt. Zusätzlich mussten die Studierenden schriftlich über Unterrichtsstunden reflektieren, um prozedurales Wissen untersuchen zu können. Diese Dokumente wurden mithilfe des Kodiermanuals auf Reflexionsbreite und Reflexionstiefe hin untersucht.

7.1.1. Ergebnisse zur Steigerung des deklarativen und konditionalen Wissens zum Thema Reflexionskompetenz (H 1)

Reflexionskompetenztest

Um den Lernerfolg der Studierenden zu messen, wurde ein Messinstrument zur Erfassung der Reflexionskompetenz zu deklarativem und konditionalem Wissen entwickelt. Die Reliabilitätsanalyse anhand der jeweiligen Post-Testdaten ergab unter Einschluss aller 17 Items ein Cronbachs Alpha von $\alpha = .60$. Werden nur Items mit einer Trennschärfe über .15 berücksichtigt, führen zwölf Items zu einer Reliabilität nach Cronbach von $\alpha = .65$.

Die eingeschränkte Reliabilität lässt sich möglicherweise darauf zurückführen, dass in die Berechnung der Reliabilität auch die Feedbackgruppe einbezogen wurde, welche individuelles externes Feedback erhielt und weniger selbst reflektieren musste.

Rückschlüsse auf die Güte des entwickelten Testinstruments lassen sich eher auf Basis der beiden selbstreflektierenden Gruppen ziehen. Werden nur die Selbstreflexions- und die Kontrollgruppe in die Reliabilitätsberechnung einbezogen, wie es inhaltlich sinnvoll erscheint, dann führen 15 Items mit einer Trennschärfe über .15 zu einer guten Reliabilität von $\alpha = .73$.

Aus Gründen der Vergleichbarkeit aller drei Gruppen werden in der Arbeit ausschließlich die Ergebnisse des auf zwölf Items gekürzten Tests mit einer Reliabilität nach Cronbach von $\alpha = .65$ berichtet. Würde man sich auf die längere Skala mit 15 Items beziehen, führte dies zu den gleichen signifikanten Unterschieden, wie sie im Folgenden aufgeführt werden. Damit kann dem Vergleichbarkeit der drei Gruppen entsprechend der aufgestellten Hypothese 6 Priorität eingeräumt werden. Zunächst werden die deskriptiven Ergebnisse und Veränderungen zwischen den Messzeitpunkten präsentiert (Abbildung 54, Tabelle 115).

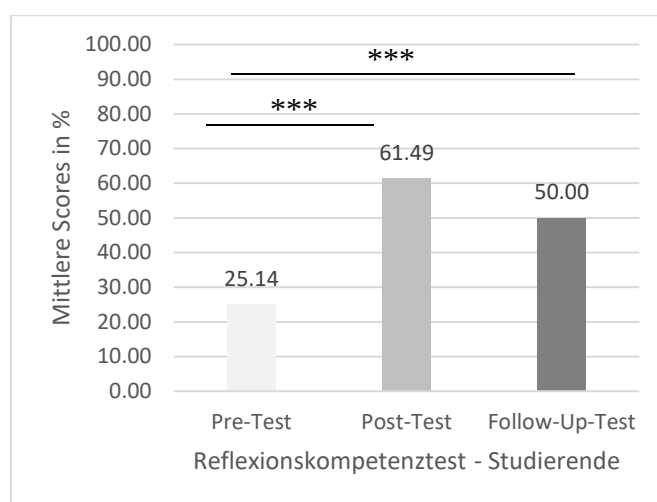


Abbildung 54: Mittlere Scores des Reflexionskompetenztests über drei Messzeitpunkte – Hauptstudie – Gesamtstichprobe

Tabelle 115: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Hauptstudie – Scores Gesamtstichprobe

Gesamtstichprobe	N	M	SD	Max.	Min.
Pre-Test	58	25.14	9.30	41.67	0.00
Post-Test	58	61.49	20.40	91.67	8.33
Follow-Up-Test	58	50.00	19.87	91.67	0.00

Wissenszuwachs:

Die weiteren Ergebnisse der Gesamtstichprobe der Hauptstudie ergeben im Hinblick auf den Reflexionskompetenztest einen nachhaltigen, höchst signifikanten Wissenszuwachs. Da nicht alle Tests normalverteilt vorlagen, werden nichtparametrische Tests berechnet. Parametrische Berechnungen können dem Anhang entnommen werden (Tabelle 226).

Wilcoxon-Test: $z_{\text{Pre-Post}} = -6.41$; $p < .001$; $r = 0.6$; $z_{\text{Pre-FU}} = -6.01$; $p < .001$; $r = 0.6$

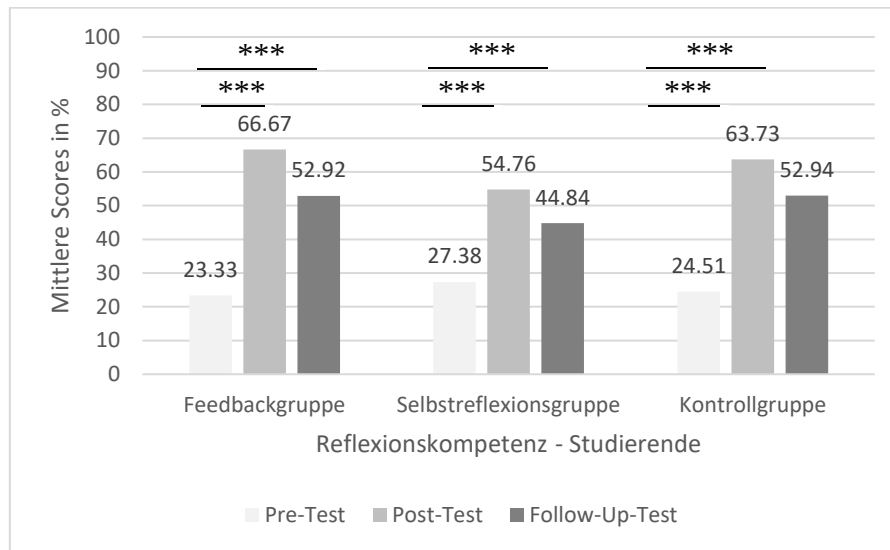
Gruppenergebnisse – deskriptiv (Abbildung 55, Tabelle 116, Tabelle 117, Tabelle 118)

Abbildung 55: Mittlere Scores des Reflexionskompetenztests über drei Messzeitpunkte – Hauptstudie – Gruppenvergleich

Tabelle 116: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Hauptstudie – Feedbackgruppe

Feedbackgruppe	N	M	SD	Max.	Min.
Pre-Test	20	23.33	7.93	41.67	16.67
Post-Test	20	66.67	15.05	91.67	33.33
Follow-Up-Test	20	52.92	18.79	83.33	16.67

Tabelle 117: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Hauptstudie – Selbstreflexionsgruppe

Selbstreflexions- gruppe	N	M	SD	Max.	Min.
Pre-Test	21	27.38	10.25	41.67	8.33
Post-Test	21	54.76	21.66	91.67	16.67
Follow-Up-Test	21	44.84	18.54	83.33	0.00

Tabelle 118: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Hauptstudie – Kontrollgruppe

Kontrollgruppe	N	M	SD	Max.	Min.
Pre-Test	17	24.51	9.53	33.33	0.00
Post-Test	17	63.73	23.00	91.67	8.33
Follow-Up-Test	17	52.94	22.43	91.67	16.67

Wissenszuwachs (Tabelle 119):

Tabelle 119: Ergebnisse des Reflexionskompetenztests – Gruppenvergleiche nichtparametrisch über insgesamt drei Messzeitpunkte – Gesamtstudie

Gruppe	Messzeitpunkte	<i>z</i>	<i>p</i>	<i>r</i>
Feedbackgruppe <i>n</i> = 20	Pre-Post	-3.93	< .001	0.6
	Pre-FU	-3.73	< .001	0.6
Selbstreflexionsgruppe <i>n</i> = 21	Pre-Post	-3.73	< .001	0.6
	Pre-FU	-3.32	= .001	0.5
Kontrollgruppe <i>n</i> = 17	Pre-Post	-3.51	< .001	0.6
	Pre-FU	-3.28	= .001	0.6

Parametrische Berechnungen können dem Anhang entnommen werden (Tabelle 227-Tabelle 229). Alle Gruppen zeigen einen nachhaltigen, höchst signifikanten Wissenszuwachs. Die höchsten parametrischen Effektstärken erreicht die Feedbackgruppe (Tabelle 230).

7.1.2. Ergebnisse zur Steigerung des prozeduralen Wissens zum Thema Reflexionskompetenz (H1 und H 2)

Kodiermanual

Das Kodiermanual wurde entwickelt, um die Steigerung des prozeduralen Wissens messen zu können. Der Einsatz erfolgte bei den schriftlichen Reflexionen nach den gehaltenen oder beobachteten Unterrichtsstunden, welche mithilfe des Manuals analysiert wurden.

Ergebnisse zur Gesamtpunktzahl – Gesamt (Abbildung 56, Tabelle 120):

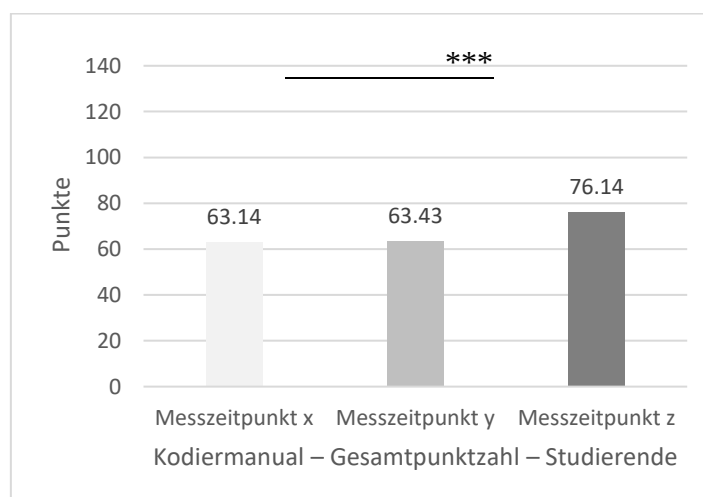


Abbildung 56: Hauptstudie – Kodiermanual – Mittlere Scores – Gesamtpunktzahl – Gesamtstichprobe – Messzeitpunkte x,y,z

Tabelle 120: Hauptstudie – Kodiermanual – Deskriptive Statistik – Gesamtpunktzahl – Gesamtstichprobe – Messzeitpunkte x,y,z

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	58	63.14	29.72	172.00	12.00
Messzeitpunkt y	58	63.43	30.51	172.00	11.00
Messzeitpunkt z	58	76.14	36.32	181.00	14.00

Da die Gesamtpunktzahl nur im Messzeitpunkt x normalverteilt ist, werden sowohl nicht-parametrische als auch parametrische Tests berechnet.

Wissenszuwachs:

Nichtparametrischer Test (Wilcoxon-Test):

$$Z_{MZPx-MZPy} = -0.20; p = .842; r = 0.02; \quad Z_{MZPx-MZPz} = -3.66; p < .001; r = 0.3;$$

Parametrischer Test: Tabelle 233

Sowohl nichtparametrische als auch parametrische Tests zeigen einen nachhaltigen, höchst signifikanten Wissenszuwachs von Messzeitpunkt x auf Messzeitpunkt z.

Ergebnisse zur Gesamtpunktzahl – Aufteilung in Gruppen (Abbildung 57, Tabelle 121, Tabelle 122, Tabelle 123):

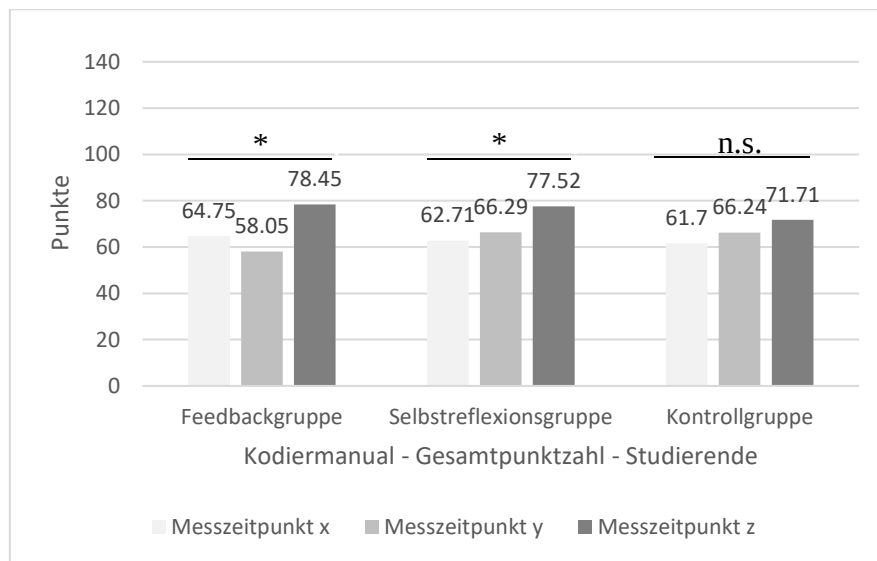


Abbildung 57: Hauptstudie – Kodiermanual – Mittlere Scores – Gesamtpunktzahl – Gruppenvergleich – Messzeitpunkte x,y,z

Tabelle 121: Hauptstudie – Kodiermanual – Deskriptive Statistik – Gesamtpunktzahl – Feedbackgruppe – Messzeitpunkte x,y,z

Feedbackgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	20	64.75	33.73	172.00	12.00
Messzeitpunkt y	20	58.05	23.28	127.00	11.00
Messzeitpunkt z	20	78.45	37.35	181.00	28.00

Tabelle 122: Hauptstudie – Kodiermanual – Deskriptive Statistik – Gesamtpunktzahl – Selbstreflexionsgruppe – Messzeitpunkte x,y,z

Selbstreflexionsgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	21	62.71	32.58	163.00	28.00
Messzeitpunkt y	21	66.29	39.89	172.00	28.00
Messzeitpunkt z	21	77.52	36.64	141.00	31.00

Tabelle 123: Hauptstudie – Kodiermanual – Deskriptive Statistik – Gesamtpunktzahl – Kontrollgruppe – Messzeitpunkte x,y,z

Kontrollgruppe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	17	61.70	21.56	93.00	33.00
Messzeitpunkt y	17	66.24	25.06	106.00	29.00
Messzeitpunkt z	17	71.71	36.53	144.00	14.00

Wissenszuwachs:

Nichtparametrischer Test (Wilcoxon-Test) (Tabelle 124):

Tabelle 124: Gesamtpunktzahl Kodiermanual – Vergleich mittlere Scores der Messzeitpunkte x,y,z – Aufteilung in Gruppen – nichtparametrisch

Gruppe	Messzeitpunkte	<i>z</i>	<i>p</i>	<i>r</i>
Feedbackgruppe <i>n</i> = 20	MZPx-MZPy	-1,55	= .121; n.s.	0.2
	MZPx-MZPz	-2.24	= .025	0.4
Selbstreflexionsgruppe <i>n</i> = 21	MZPx-MZPy	-0.78	= .433; n.s.	0.1
	MZPx-MZPz	-2.40	= .016	0.4
Kontrollgruppe <i>n</i> = 17	MZPx-MZPy	-1.14	= .255; n.s.	0.2
	MZPx-MZPz	-1.42	= .155; n.s.	0.2

Parametrischer Test: Tabelle 237 bis Tabelle 239

Von MZPx auf MZPy ist bei allen Gruppen keine signifikante Steigerung erkennbar. Der Vergleich von MZPx und MZPz lässt in den Interventionsgruppen eine signifikante Steigerung erkennen.

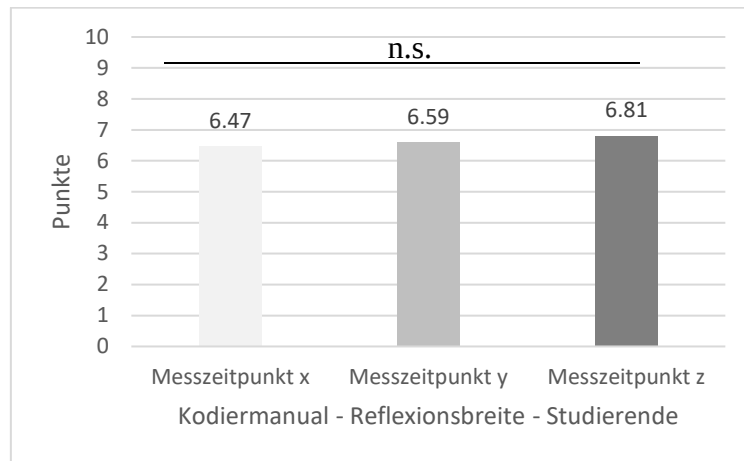
Ergebnisse zur Reflexionsbreite – Gesamt (Abbildung 58, Tabelle 125):

Abbildung 58: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Gesamtstichprobe – Messzeitpunkte x,y,z

Tabelle 125: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Gesamtstichprobe – Messzeitpunkte x,y,z

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	58	6.47	1.65	10.00	3.00
Messzeitpunkt y	58	6.59	1.49	10.00	2.00
Messzeitpunkt z	58	6.81	1.88	9.00	1.00

Da der Test auf Normalverteilung signifikante Werte ergibt, werden im Folgenden nichtparametrische Tests berechnet.

Wissenszuwachs:

Nichtparametrischer Test (Wilcoxon-Test):

$$Z_{MZPx-MZPy} = -.73; p = .466; \text{n.s.}; r = 0.1;$$

$$Z_{MZPx-MZPz} = -1.47; p = .140; \text{n.s.}; r = 0.1;$$

$$Z_{MZPy-MZPz} = -1.02; p = .308; \text{n.s.}; r = 0.1;$$

In der Reflexionsbreite konnten sich die Gruppen gesamt gesehen nicht signifikant steigern.

Ergebnisse zur Reflexionsbreite – Aufteilung in Gruppen (Abbildung 59, Tabelle 126, Tabelle 127, Tabelle 128):

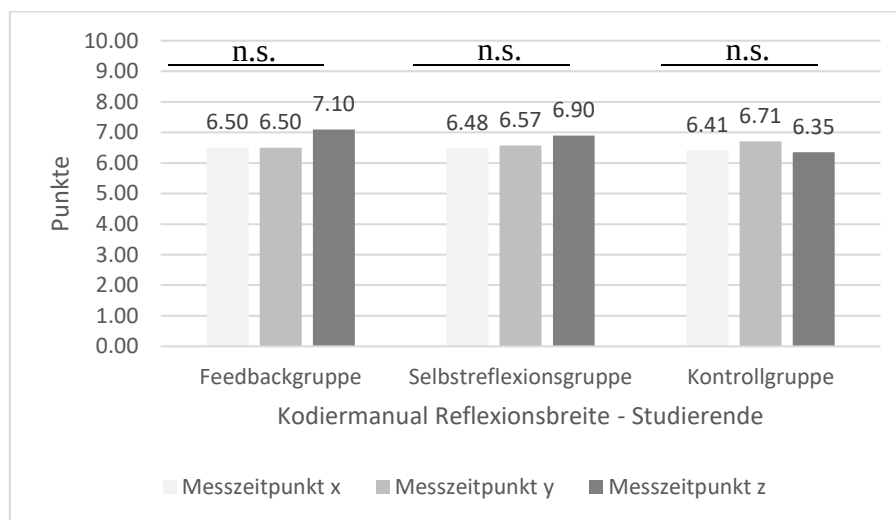


Abbildung 59: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Aufteilung in Gruppen – Messzeitpunkte x,y,z

Tabelle 126: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Feedbackgruppe – Messzeitpunkte x,y,z

Feedbackgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	20	6.50	1.61	9.00	3.00
Messzeitpunkt y	20	6.50	1.54	8.00	2.00
Messzeitpunkt z	20	7.10	1.68	9.00	3.00

Tabelle 127: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Selbstreflexionsgruppe – Messzeitpunkte x,y,z

Selbstreflexionsgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	21	6.48	1.63	10.00	3.00
Messzeitpunkt y	21	6.57	1.47	10.00	4.00
Messzeitpunkt z	21	6.90	1.97	9.00	3.00

Tabelle 128: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Kontrollgruppe – Messzeitpunkte x,y,z

Kontrollgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	17	6.41	1.80	9.00	3.00
Messzeitpunkt y	17	6.71	1.53	9.00	4.00
Messzeitpunkt z	17	6.35	2.00	9.00	1.00

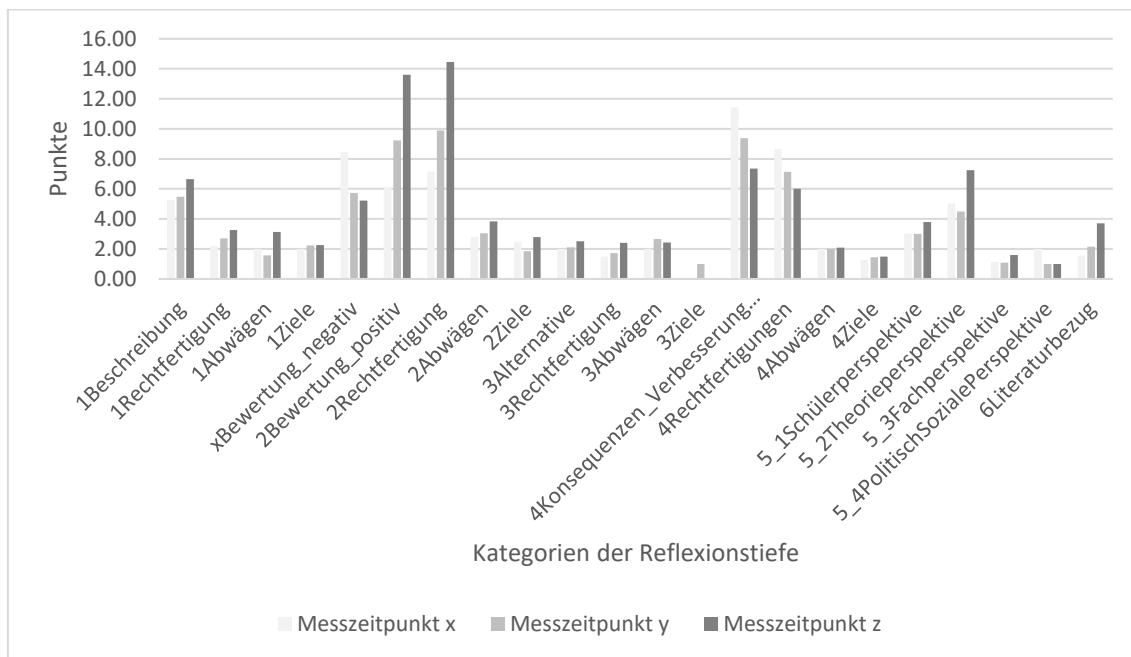
Wissenszuwachs (Tabelle 129):*Tabelle 129: Wissenszuwachs Reflexionsbreite – Aufteilung in Gruppen und Messzeitpunkte x,y,z*

Gruppe	Messzeitpunkte	<i>z</i>	<i>p</i>	<i>r</i>
Feedbackgruppe <i>n</i> = 20	MZPx – MZPy	-0.10	= .921; n.s	0.02
	MZPx – MZPz	-1.61	= .106; n.s	0.3
	MZPy – MZPz	-1.57	= .116; n.s	0.2
Selbstreflexionsgruppe <i>n</i> = 21	MZPx – MZPy	-0.44	= .660; n.s	0.1
	MZPx – MZPz	-1.08	= .280; n.s	0.2
	MZPy – MZPz	-0.94	= .345; n.s	0.1
Kontrollgruppe <i>n</i> = 17	MZPx – MZPy	-0.75	= .453; n.s	0.1
	MZPx – MZPz	-0.03	= .979; n.s	0
	MZPy – MZPz	-0.97	= .334; n.s	0.2

Es ist in keiner der Gruppen ein signifikanter Wissenszuwachs zu erkennen.

Ergebnisse zur Reflexionstiefe – Gesamt:

Entwicklung der einzelnen Kategorien – Reflexionstiefe – deskriptiv (Abbildung 60):

*Abbildung 60: Hauptstudie – Mittlere Scores der Kategorien der Reflexionstiefe über drei Messzeitpunkte x,y,z*

Die Tabelle mit den zugehörigen Werten befindet sich im Anhang (Tabelle 245). Es folgen die Berechnungen zur Reflexionstiefe der Gesamtstichprobe (Abbildung 61, Tabelle 130):

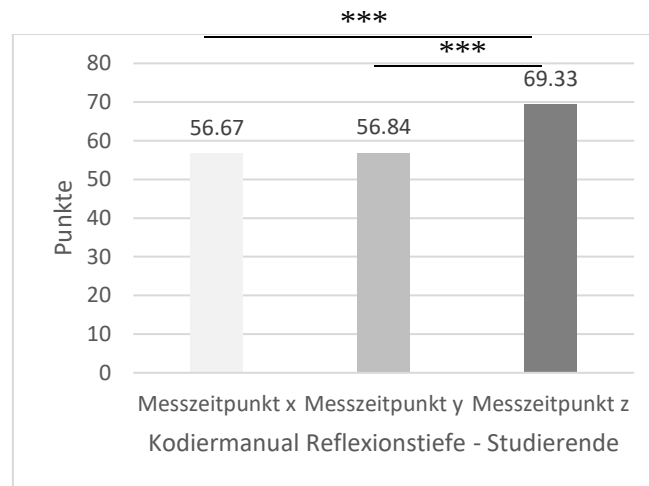


Abbildung 61: Hauptstudie – Kodiermanual – Mittlere Scores – Summe der Reflexionstiefe – Gesamtstichprobe – Messzeitpunkte x,y,z

Tabelle 130: Hauptstudie – Kodiermanual – Deskriptive Statistik – Summe der Reflexionstiefe – Gesamtstichprobe – Messzeitpunkte x,y,z

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	58	56.67	28.85	163.00	9.00
Messzeitpunkt y	58	56.84	29.54	162.00	9.00
Messzeitpunkt z	58	69.33	34.97	172.00	13.00

Da die drei Tests nicht alle normalverteilt vorliegen, werden sowohl nichtparametrische als auch parametrische Tests berechnet.

Wissenszuwachs:

Nichtparametrischer Test (Wilcoxon-Test):

$$z_{MZPx-MZPy} = -0.18; p = .858; \text{n.s.}; r = 0.02;$$

$$z_{MZPx-MZPz} = -3.74; p < .001; r = 0.3;$$

$$z_{MZPy-MZPz} = -3.42; p = .001; r = 0.3;$$

Parametrischer Test: Tabelle 248

In der Reflexionstiefe zeigt sich in der Gesamtstichprobe ein höchst signifikanter Lernzuwachs von MZPx auf MZPz und von MZPy auf MZPz.

Ergebnisse zur Reflexionstiefe – Aufteilung in Gruppen (Abbildung 62, Tabelle 131, Tabelle 132, Tabelle 133):

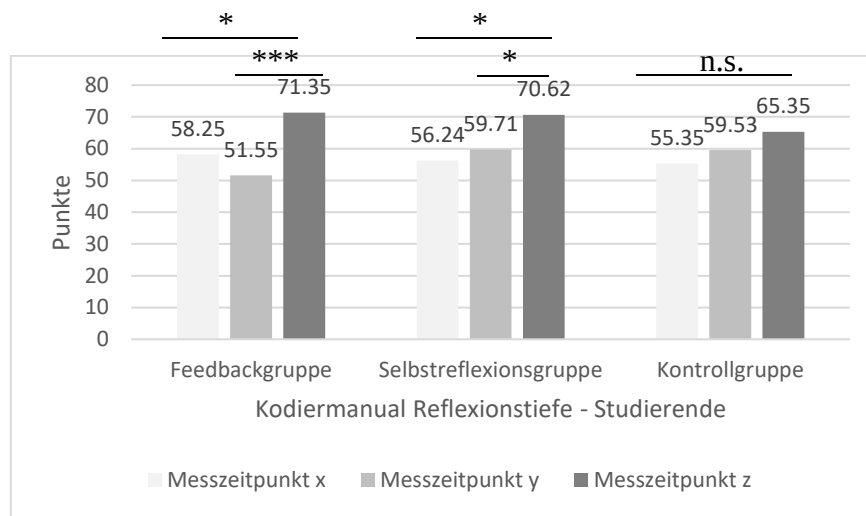


Abbildung 62: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionstiefe – Gruppenvergleiche – Messzeitpunkte x,y,z

Tabelle 131: Hauptstudie – Kodiermanual – Deskriptive Statistik – Reflexionstiefe – Feedbackgruppe – Messzeitpunkte x,y,z

Feedbackgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	20	58.25	32.59	163.00	9.00
Messzeitpunkt y	20	51.55	22.47	119.00	9.00
Messzeitpunkt z	20	71.35	36.14	172.00	25.00

Tabelle 132: Hauptstudie – Kodiermanual – Deskriptive Statistik – Reflexionstiefe – Selbstreflexionsgruppe – Messzeitpunkte x,y,z

Selbstreflexionsgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	21	56.24	31.61	153.00	23.00
Messzeitpunkt y	21	59.71	38.76	162.00	22.00
Messzeitpunkt z	21	70.62	35.06	132.00	25.00

Tabelle 133: Hauptstudie – Kodiermanual – Deskriptive Statistik – Reflexionstiefe – Kontrollgruppe – Messzeitpunkte x,y,z

Kontrollgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	17	55.35	21.18	88.00	24.00
Messzeitpunkt y	17	59.53	24.03	99.00	22.00
Messzeitpunkt z	17	65.35	35.29	136.00	13.00

Wissenszuwachs:

Nichtparametrische Tests (Wilcoxon-Test) (Tabelle 134):

Tabelle 134: Wissenszuwachs Reflexionstiefe – Aufteilung in Gruppen und Messzeitpunkte x,y,z

Gruppe	Messzeitpunkte	z	p	r
Feedbackgruppe n = 20	MZPx-MZPy	-1.44	= .150; n.s.	0.2
	MZPx-MZPz	-2.28	= .023	0.4
	MZPy-MZPz	-3.25	=.001	0.5
Selbstreflexionsgruppe n = 21	MZPx-MZPy	-0.73	= .465; n.s.	0.1
	MZPx-MZPz	-2.47	= .014	0.4
	MZPy-MZPz	-1.97	= .049	0.3
Kontrollgruppe n = 17	MZPx-MZPy	-1.09	= .276; n.s.	0.2
	MZPx-MZPz	-1.49	= .136; n.s.	0.3
	MZPy-MZPz	-0.90	= .368; n.s.	0.2

Parametrische Tests: Tabelle 250, Tabelle 252 und Tabelle 254

7.2. Ergebnisse der Zusammenhänge der Förderung der Reflexionskompetenz (Forschungsfrage 2)

Im Theorieteil wurde bereits von positiven Korrelationen der Reflexionskompetenz mit verschiedenen Bereichen berichtet. Zunächst folgen die Ergebnisse des fachdidaktischen Wissens und des Fachwissens, im Anschluss wird der residuale Lernzuwachs der beiden Wissensbereiche mit dem residualen Lernzuwachs im Bereich Reflexionskompetenz korreliert. Dies erfolgt zunächst mit der Gesamtstichprobe. Ebenso wird bezüglich Zielorientierungen, fachspezifischen Überzeugungen und Selbstwirksamkeitserwartungen verfahren. Zur Prüfung von Hypothese 5 werden die Leistungen der Schülerinnen und Schüler untersucht und diese in Korrelation mit den Ergebnissen der Studierenden gesetzt. Für Hypothese 6 werden die residualen Lernzuwächse der einzelnen Gruppen in den Bereichen Reflexionskompetenz, fachdidaktisches Wissen, Fachwissen, Selbstwirksamkeitserwartung und Unterrichtsqualität verglichen.

7.2.1. Ergebnisse der Bereiche fachdidaktisches Wissen und Fachwissen sowie der Korrelationen bezüglich der Reflexionskompetenz mit den Bereichen fachdidaktisches Wissen und Fachwissen (H3a und H3b)

Test zum fachdidaktischen Wissen

Wie in Kapitel 6.1.1 berichtet, wurden in diesem Test 38 Paarvergleiche gezogen und eine zufriedenstellende Reliabilität mit $\alpha = .85$ erzielt (38 Relationen; maximal zu erreichende Punkte: 38). Die mittleren Scores der Gesamtstichprobe verteilen sich wie folgt (Abbildung 63, Tabelle 135):

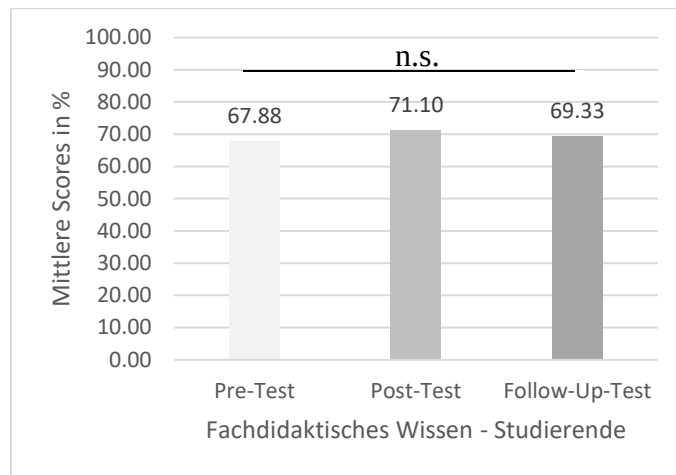


Abbildung 63: Hauptstudie – Mittlere Scores – fachdidaktisches Wissen – Scores Gesamtstichprobe über drei Messzeitpunkte

Tabelle 135: Hauptstudie – Deskriptive Statistik – fachdidaktisches Wissen – Scores Gesamtstichprobe über drei Messzeitpunkte

Fachdidaktisches Wissen	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	58	67.88	12.96	97.37	36.84
Post-Test	58	71.10	16.17	94.74	21.05
Follow-Up-Test	58	69.33	11.99	94.74	39.47

Wissenszuwachs:

Da lediglich der Follow-Up-Test normalverteilt vorliegt, werden nichtparametrische Tests berechnet. Zum Vergleich finden sich die parametrischen Tests im Anhang (Tabelle 257).

Nichtparametrischer Test (Wilcoxon-Test):

$$z_{\text{Pre-Post}} = -1.82; p = .069; \text{n.s.}; r = 0.2;$$

$$z_{\text{Pre-FU}} = -0.83; p = .405; \text{n.s.}; r = 0.1;$$

Der Vergleich von Pre- und Post-Test ist knapp nicht signifikant. Ein nachhaltiger Lerneffekt im Bereich fachdidaktisches Wissen ist nicht zu verzeichnen.

Test zum Fachwissen

Der Test erreicht eine zufriedenstellende Reliabilität mit $\alpha = .72$ bei 19 Items. Die mittleren Scores der Gesamtstichprobe verteilen sich wie folgt (Abbildung 64, Tabelle 136):

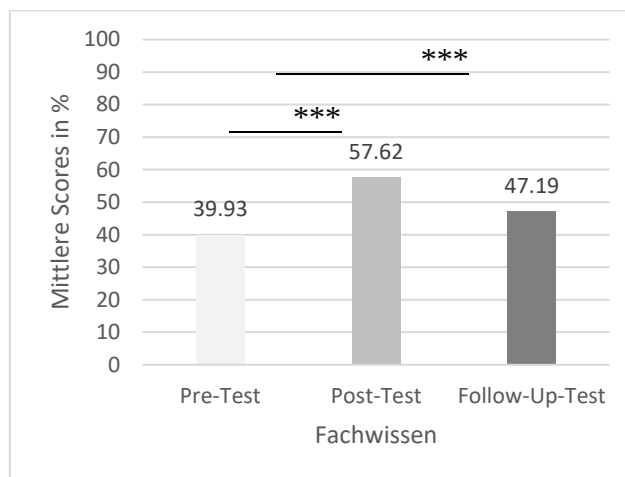


Abbildung 64: Hauptstudie – Mittlere Scores – Fachwissen – Scores Gesamtstichprobe über drei Messzeitpunkte

Tabelle 136: Hauptstudie – Deskriptive Statistik – Fachwissen – Scores Gesamtstichprobe über drei Messzeitpunkte

Fachwissen	N	M	SD	Max.	Min.
Pre-Test	58	39.93	13.80	63.16	0.00
Post-Test	58	57.62	19.13	94.74	15.79
Follow-Up-Test	58	47.19	15.40	84.21	21.05

Wissenszuwachs:

Da der Pre-Test nicht normalverteilt vorliegt, werden nichtparametrische Tests berechnet.

Nichtparametrischer Test (Wilcoxon-Test):

$$Z_{\text{Pre-Post}} = -5.54; p < .001; r = 0.5;$$

$$Z_{\text{Pre-FU}} = -3.45; p = .001; r = 0.3;$$

Parametrischer Test (Tabelle 260)

Der Lernzuwachs sowohl von Pre- auf Post-Test als auch Pre- auf Follow-Up-Test ist hoch signifikant.

Korrelationen des residualen Lernzuwachses:

Korrelationen der Gesamtstichprobe

Für normalverteilte Datenpaare werden die Korrelationen nach Pearson berechnet. Nicht normalverteilte Korrelationen werden nach Spearman angegeben. Bei Kombinationen aus normalverteilten und nicht normalverteilten Daten werden beide Ergebnisse berichtet (Tabelle 137).

Tabelle 137: Hauptstudie – Korrelationen des residualen Lernzuwachses – fachdidaktisches Wissen und Fachwissen mit Kodiermanual – Gesamtstichprobe

Residualer Lernzuwachs	Fachdidaktisches Wissen	Fachwissen
Kodiermanual Reflexionsbreite	$r(58) = -.08; p = .560; \text{n.s.}$ $r_s(58) = -.11; p = .427; \text{n.s.}$	$r(58) = -.03; p = .848; \text{n.s.}$ $r_s(58) = -.08; p = .563; \text{n.s.}$
Kodiermanual Reflexionstiefe	$r_s(58) = -.07; p = .599.; \text{n.s.}$	$r_s(58) = -.03; p = .820; \text{n.s.}$
Kodiermanual Perspektiven	$r(58) = .002; p = .986; \text{n.s.}$ $r_s(58) = .06; p = .655; \text{n.s.}$	$r(58) = .09; p = .486; \text{n.s.}$ $r_s(58) = .09; p = .485; \text{n.s.}$
Kodiermanual Literaturbezug	$r_s(58) = -.03; p = .828; \text{n.s.}$	$r_s(58) = .04; p = .796; \text{n.s.}$
Kodiermanual Gesamtpunktzahl	$r_s(58) = -.05; p = .693; \text{n.s.}$	$r_s(58) = .04; p = .778; \text{n.s.}$
Reflexionskompetenztest	$r(58) = .18; p = .182; \text{n.s.}$ $r_s(58) = .21; p = .107; \text{n.s.}$	$r(58) = .18; p = .171; \text{n.s.}$ $r_s(58) = .19; p = .163; \text{n.s.}$

Es können keine Zusammenhänge des fachdidaktischen Wissens und des Fachwissens mit den verschiedenen Wissensbereichen der Reflexionskompetenz beobachtet werden.

7.2.2. Ergebnisse der Korrelationen bezüglich der Reflexionskompetenz und dem Studierendenfragebogen (H 4)

Bei dem Studierendenfragebogen ergeben sich folgende Reliabilitäten (Tabelle 138):

Tabelle 138: Hauptstudie – Studierendenfragebogen – Reliabilitäten – Gesamtstichprobe

Unterpunkt	Cronbachs Alpha	Anzahl der Items
Zielorientierungen	$\alpha = .67$	16 Items
Fachspezifische Überzeugungen bzgl. Chemie als Wissenschaft	$\alpha = .53$	10 Items
Fachspezifische Überzeugungen bzgl. Lehren von Chemie	$\alpha = .78$	21 Items
Fachspezifische Überzeugungen bzgl. Lernen von Chemie	$\alpha = .50$	8 Items
Selbstwirksamkeitserwartung bzgl. der Planung von Experimenten	$\alpha = .67$	5 Items
Selbstwirksamkeitserwartung bzgl. der Durchführung von Experimenten	$\alpha = .74$	4 Items
Selbstwirksamkeitserwartung bzgl. der didaktischen Rekonstruktion	$\alpha = .77$	8 Items
Selbstwirksamkeitserwartung bzgl. der Umsetzung der didaktischen Rekonstruktion	$\alpha = .69$	5 Items

Obwohl die Reliabilitätswerte erst ab $\alpha = .70$ als akzeptabel gelten, werden niedrigere Werte als Kontrollvariable verwendet (Rost, 2013). Diese werden lediglich als Tendenzen ausgelegt.

Zielorientierungen

Die mittleren Scores der Gesamtstichprobe verteilen sich wie folgt (Abbildung 65, Tabelle 139):

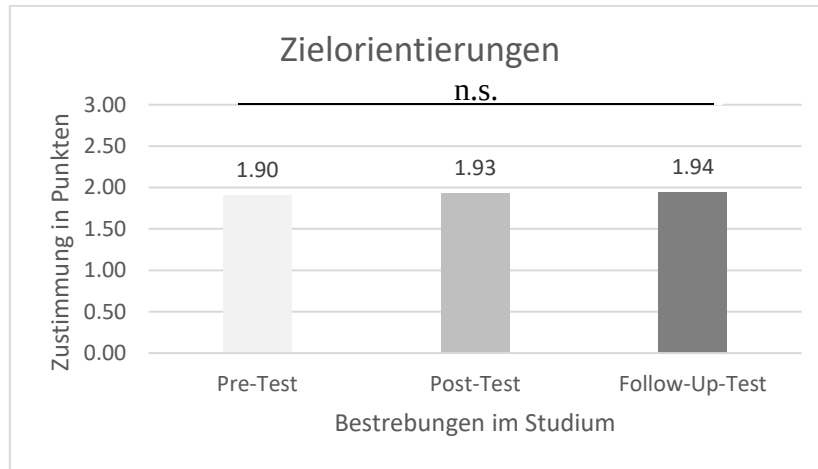


Abbildung 65: Hauptstudie – Mittlere Scores – Studierendenfragebogen – Zielorientierungen – Gesamtstichprobe über drei Messzeitpunkte

Tabelle 139: Hauptstudie – Deskriptive Statistik – Studierendenfragebogen – Zielorientierungen – Gesamtstichprobe über drei Messzeitpunkte

Bestrebungen im Studium	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Pre-Test	58	1.90	0.35	2.69	1.13
Post-Test	58	1.93	0.32	2.56	1.06
Follow-Up-Test	58	1.94	0.34	2.88	1.13

Entwicklung über drei Messzeitpunkte:

Da alle Messzeitpunkt normalverteilt vorliegen, wird der *t*-Test für verbundene Stichproben als parametrischer Test berechnet.

Pre – Post: $t(57) = -0.665$; $p = .509$; n.s.; $d = 0.1$;

Pre – FU: $t(57) = -1.010$; $p = .317$; n.s.; $d = 0.1$;

Es gibt keine signifikante Entwicklung im Bereich der Zielorientierungen.

Fachspezifische Überzeugungen

Die mittleren Scores der Gesamtstichprobe verteilen sich wie folgt (Abbildung 66, Tabelle 140):

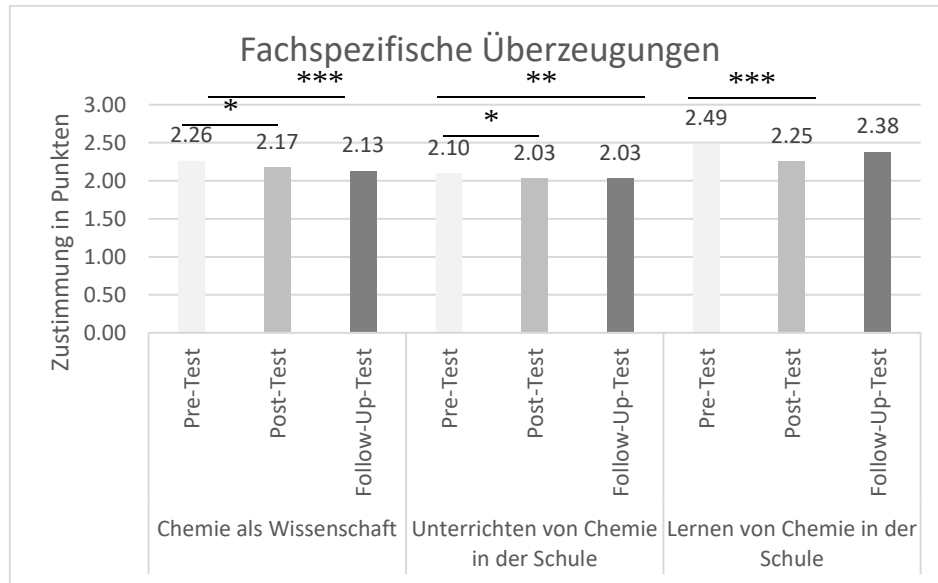


Abbildung 66: Hauptstudie – Mittlere Scores – Studierendenfragebogen – Fachspezifische Überzeugungen – Gesamtstichprobe über drei Messzeitpunkte

Tabelle 140: Hauptstudie – Deskriptive Statistik – Studierendenfragebogen – Fachspezifische Überzeugungen – Gesamtstichprobe über drei Messzeitpunkte

		<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Chemie als Wissenschaft	Pre-Test	58	2.26	0.37	3.00	1.50
	Post-Test	58	2.17	0.37	3.20	1.40
	Follow-Up-Test	58	2.13	0.35	2.90	1.30
Unterrichten von Chemie in der Schule	Pre-Test	58	2.10	0.28	2.76	1.57
	Post-Test	58	2.03	0.31	2.86	1.10
	Follow-Up-Test	58	2.03	0.32	3.00	1.24
Lernen von Chemie in der Schule	Pre-Test	58	2.49	0.36	3.25	1.75
	Post-Test	58	2.25	0.35	3.50	1.38
	Follow-Up-Test	58	2.38	0.34	3.25	1.63

Entwicklung über drei Messzeitpunkte:

Da die Messzeitpunkte teilweise normalverteilt und teilweise nicht normalverteilt vorliegen, werden nichtparametrische Tests berechnet. Parametrische Tests können dem Anhang entnommen werden (Tabelle 273).

Nichtparametrischer Test (Wilcoxon-Test):

Chemie als Wissenschaft:

$z_{\text{Pre-Post}} = -1.97; p = .048; r = 0.2;$

$z_{\text{Pre-FU}} = -3.59; p < .001; r = 0.3;$

Unterrichten von Chemie in der Schule:

$z_{\text{Pre-Post}} = -2.14; p = .032; r = 0.2;$

$z_{\text{Pre-FU}} = -2.82; p = .005; r = 0.3;$

Lernen von Chemie in der Schule:

$z_{\text{Pre-Post}} = -4.04; p < .001; r = 0.4;$

$z_{\text{Pre-FU}} = -1.71; p = .087; \text{n.s.}; r = 0.2;$

Die Teilbereiche Chemie als Wissenschaft und Unterrichten von Chemie in der Schule konnten sich signifikant verbessern. Der Bereich Lernen von Chemie in der Schule weist nur vom ersten auf den zweiten Messzeitpunkt eine höchst signifikante Steigerung auf. Im Vergleich vom ersten auf den dritten Messzeitpunkt ist keine signifikante Steigerung zu beobachten.

Selbstwirksamkeitserwartungen

Die mittleren Scores der Gesamtstichprobe verteilen sich wie folgt (Abbildung 67, Tabelle 141):

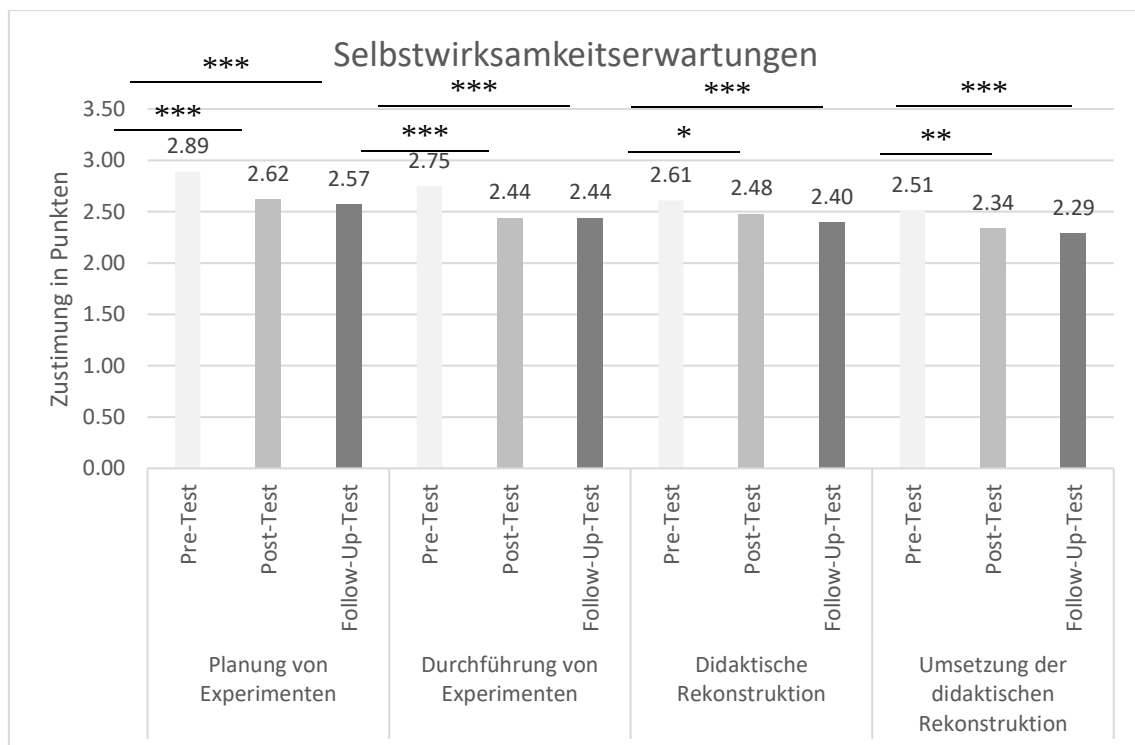


Abbildung 67: Hauptstudie – Mittlere Scores – Studierendenfragebogen – Selbstwirksamkeitserwartungen – Gesamtstichprobe über drei Messzeitpunkte

Tabelle 141: Hauptstudie – Deskriptive Statistik – Studierendenfragebogen – Selbstwirksamkeitserwartungen – Gesamtstichprobe über drei Messzeitpunkte

		<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Planung von Experimenten	Pre-Test	58	2.89	0.49	4.20	1.80
	Post-Test	57	2.62	0.51	4.00	1.60
	Follow-Up-Test	58	2.57	0.43	3.60	1.40
Durchführung von Experimenten	Pre-Test	58	2.75	0.49	3.75	1.75
	Post-Test	57	2.44	0.51	3.50	1.00
	Follow-Up-Test	58	2.44	0.45	3.25	1.25
Didaktische Rekonstruktion unterrichtsrelevanter Inhalte	Pre-Test	58	2.61	0.42	4.00	1.75
	Post-Test	58	2.48	0.43	3.88	1.50
	Follow-Up-Test	58	2.40	0.43	3.50	1.38
Umsetzung der didaktischen Rekonstruktion im Unterricht	Pre-Test	58	2.51	0.55	4.00	1.40
	Post-Test	58	2.34	0.46	3.20	1.40
	Follow-Up-Test	58	2.29	0.43	3.40	1.20

Entwicklung über drei Messzeitpunkte:

Da die Messzeitpunkte teilweise normalverteilt und teilweise nicht normalverteilt vorliegen, werden nichtparametrische Tests berechnet. Parametrische Tests können dem Anhang entnommen werden (Tabelle 275):

Nichtparametrischer Test (Wilcoxon-Test):

Planung von Experimenten:

$$z_{\text{Pre-Post}} = -3.41; p = .001; r = 0.3;$$

$$z_{\text{Pre-FU}} = -4.77; p < .001; r = 0.4;$$

Durchführung von Experimenten:

$$z_{\text{Pre-Post}} = -4.15; p < .001; r = 0.4;$$

$$z_{\text{Pre-FU}} = -3.83; p < .001; r = 0.4;$$

Didaktische Rekonstruktion:

$$z_{\text{Pre-Post}} = -2.33; p = .020; r = 0.2;$$

$$z_{\text{Pre-FU}} = -3.67; p < .001; r = 0.3;$$

Umsetzung der didaktischen Rekonstruktion:

$$z_{\text{Pre-Post}} = -2.62; p = .009; r = 0.2;$$

$$z_{\text{Pre-FU}} = -3.53; p < .001; r = 0.3;$$

Alle gemessenen Bereiche der Selbstwirksamkeitserwartung verbessern sich über die drei Messzeitpunkte höchst signifikant.

Korrelationen des residualen Lernzuwachses der Gesamtstichprobe:

Da die residualen Lernzuwächse nicht immer normalverteilt vorliegen, werden dementsprechend für normalverteilte Datenpaare die Korrelationen nach Pearson berechnet. Nicht normalverteilte Korrelationen werden nach Spearman angegeben. Bei Kombinationen aus normalverteilten und nicht normalverteilten Daten werden beide Ergebnisse berichtet (Tabelle 142).

Tabelle 142: Korrelationen des residualen Lernzuwachses der Gesamtstichprobe

Residualer Lernzuwachs	Zielorientierungen	Fachspezifische Überzeugungen Chemie als Wissenschaft	Fachspezifische Überzeugungen Unterrichten von Chemie	Fachspezifische Überzeugungen Lernen von Chemie
Kodiermanual Reflexionsbreite	$r(58) = -.25$; $p = .054$; n.s.	$r(58) = -.35$; $p = .007$	$r(58) = -.16$; $p = .219$; n.s.	$r(58) = -.15$; $p = .267$; n.s.
Kodiermanual Reflexionstiefe	$r(58) = -.23$; $p = .077$; n.s. $r_s(58) = -.25$; $p = .056$; n.s.	$r(58) = -.22$; $p = .092$; n.s. $r_s(58) = -.19$; $p = .160$; n.s.	$r(58) = -.34$; $p = .010$ $r_s(58) = -.34$; $p = .009$	$r(58) = -.20$; $p = .141$; n.s. $r_s(58) = -.21$; $p = .110$; n.s.
Kodiermanual Gesamtpunktzahl	$r(58) = -.24$; $p = .067$; n.s. $r_s(58) = -.26$; $p = .047$	$r(58) = -.24$; $p = .065$; n.s. $r_s(58) = -.22$; $p = .092$; n.s.	$r(58) = -.34$; $p = .009$ $r_s(58) = -.35$; $p = .007$	$r(58) = -.20$; $p = .125$; n.s. $r_s(58) = -.25$; $p = .056$; n.s.
Reflexionskompetenztest	$r(58) = .12$; $p = .385$	$r(58) = .14$; $p = .291$; n.s.	$r(58) = .05$; $p = .694$; n.s.	$r(58) = .30$; $p = .021$

Residualer Lernzuwachs	Selbstwirksamkeitserwartungen Planung von Experimenten	Selbstwirksamkeitserwartungen Durchführung von Experimenten	Selbstwirksamkeitserwartungen Didaktische Rekonstruktion	Selbstwirksamkeitserwartungen Umsetzung der didaktischen Rekonstruktion
Kodiermanual Reflexionsbreite	$r(58) = -.19$; $p = .164$; n.s.	$r(58) = -.14$; $p = .312$; n.s.	$r(58) = -.06$; $p = .654$; n.s.	$r(58) = -.12$; $p = .367$; n.s.
Kodiermanual Reflexionstiefe	$r(58) = -.09$; $p = .512$; n.s. $r_s(58) = -.01$; $p = .973$; n.s.	$r(58) = -.18$; $p = .168$; n.s. $r_s(58) = -.17$; $p = .206$; n.s.	$r(58) = -.18$; $p = .177$; n.s. $r_s(58) = -.24$; $p = .067$; n.s.	$r(58) = -.17$; $p = .207$; n.s. $r_s(58) = -.14$; $p = .301$; n.s.
Kodiermanual Gesamtpunktzahl	$r(58) = -.09$; $p = .497$; n.s. $r_s(58) = -.001$; $p = .995$; n.s.	$r(58) = -.18$; $p = .178$; n.s. $r_s(58) = -.15$; $p = .248$; n.s.	$r(58) = -.19$; $p = .151$; n.s. $r_s(58) = -.25$; $p = .054$; n.s.	$r(58) = -.17$; $p = .216$; n.s. $r_s(58) = -.14$; $p = .286$; n.s.

Reflexionskompetenztest	$r(58) = .14$; $p = .289$	$r(58) = .18$; $p = .186$; n.s.	$r(58) = .01$; $p = .942$; n.s.	$r(58) = .31$; $p = .017$

Die Zielorientierungen korrelieren im nichtparametrischen Test mit der Gesamtpunktzahl des Kodiermanuals. Die drei Unterpunkte der fachspezifischen Überzeugungen korrelieren unterschiedlich: Chemie als Wissenschaft korreliert mit der Reflexionsbreite, Unterrichten von Chemie sowohl mit der Reflexionstiefe als auch mit der Gesamtpunktzahl und Lernen von Chemie korreliert mit dem Reflexionskompetenztest. Bei den Selbstwirksamkeitserwartungen korreliert lediglich der Unterpunkt Umsetzung der didaktischen Konstruktion mit dem Reflexionskompetenztest.

7.2.3. Ergebnisse des Wissenstests der Schülerinnen und Schüler über drei Messzeitpunkte (H 5)

Im Folgenden werden die deskriptiven Ergebnisse des Wissenstests der Schülerinnen und Schüler präsentiert.

Ergebnisse – Gesamt (Abbildung 68, Tabelle 143):

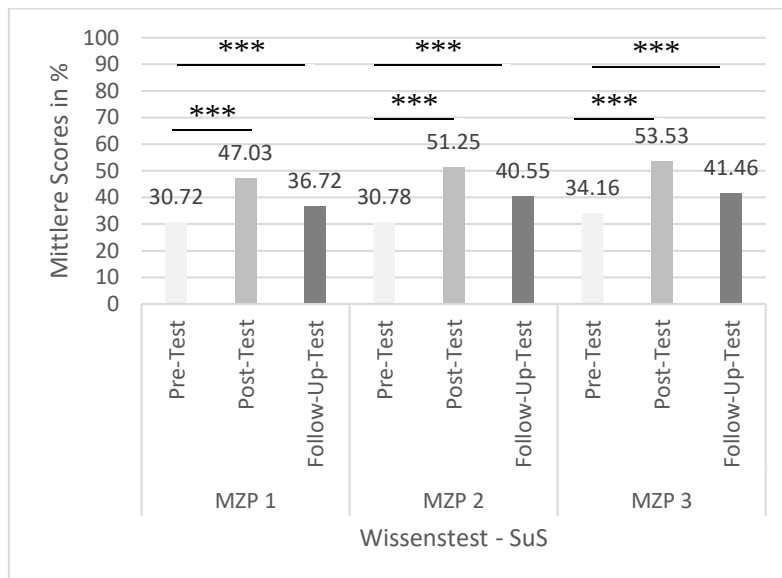


Abbildung 68: Hauptstudie – Mittlere Scores über drei MZP – Wissenstest Schülerinnen und Schüler

Tabelle 143: Hauptstudie – Deskriptive Statistik über drei MZP – Scores Wissenstest Schülerinnen und Schüler

		<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
MZP 1	Pre-Test	170	30.72	10.55	58.82	3.33
	Post-Test	170	47.03	14.01	88.24	13.33
	Follow-Up-Test	170	36.72	14.98	78.57	7.14
MZP 2	Pre-Test	171	30.78	12.85	70.59	3.33
	Post-Test	171	51.25	14.35	94.12	16.67
	Follow-Up-Test	171	40.55	14.85	82.35	7.14
MZP 3	Pre-Test	159	34.16	11.59	64.71	6.67

Post-Test	159	53.53	16.14	88.24	17.65
Follow-Up-Test	159	41.46	14.18	76.47	11.76

Da die Tests nicht immer normalverteilt sind, werden nichtparametrische Mittelwertvergleiche berechnet. Die parametrischen Ergebnisse finden sich im Anhang (Tabelle 285 bis Tabelle 287):

Nichtparametrischer Test (Wilcoxon-Test) (Tabelle 144):

Tabelle 144: Mittelwertvergleiche - Wissenstest Schülerinnen und Schüler über drei Messzeitpunkte

Messzeitpunkt Stichprobe		<i>z</i>	<i>p</i>	<i>r</i>
MZP 1 <i>n</i> = 170	Pre-Post	-9.51	< .001	0.5
	Pre-FU	-4.58	< .001	0.2
MZP 2 <i>n</i> = 171	Pre-Post	-10.33	< .001	0.6
	Pre-FU	-7.07	< .001	0.4
MZP 3 <i>n</i> = 159	Pre-Post	-9.88	< .001	0.6
	Pre-FU	-5.58	< .001	0.3

Vergleich des Z-standardisierten residualen Lernzuwachses über die drei Messzeitpunkte (Abbildung 69, Tabelle 145):

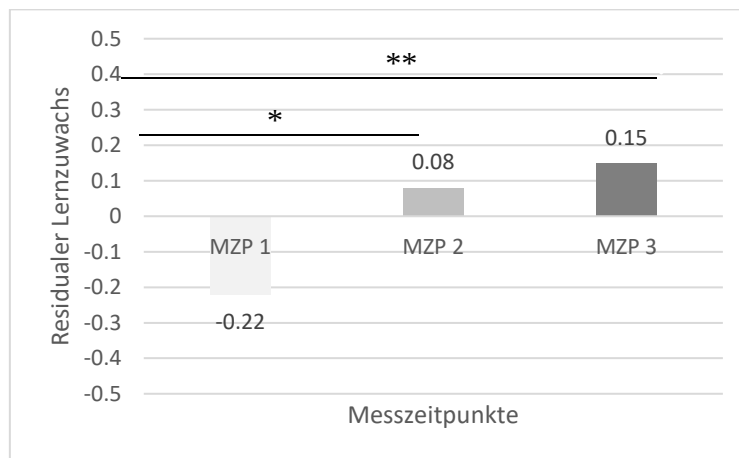


Abbildung 69: Hauptstudie – Residualer Lernzuwachs über drei Messzeitpunkte – Schülerinnen und Schüler

Tabelle 145: Hauptstudie – Deskriptive Statistik – Residualer Lernzuwachs über drei Messzeitpunkte – Schülerinnen und Schüler

Residualer Lernzuwachs	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
MZP 1	170	-.22	0.96	2.67	-2.67
MZP 2	171	.08	0.95	2.82	-2.07
MZP 3	159	.15	1.06	2.82	-2.26

Es ist zu beobachten, dass die Schülerinnen und Schüler aus dem Messzeitpunkt 1 hinter den mittleren Erwartungen zurückbleiben. Im Messzeitpunkt 2 werden leicht höhere Ergebnisse als erwartet erzielt, diese steigern sich zu Messzeitpunkt 3. Insgesamt beträgt die residuale Differenz zwischen MZP 1 und 2 über eine Viertel Standardabweichung, während diese zwischen MZP 2 und 3 mit 0.07 nicht auf vergleichbarem Niveau liegt. Die einfaktorielle Varianzanalyse lässt Unterschiede zwischen den drei Messzeitpunkten nachweisen ($F(2,499) = 6.32$; $p = .002$). Der Post-Hoc-Test belegt diese Unterschiede zwischen MZP 1 und MZP 2 ($p = .020$) und zwischen MZP 1 und MZP 3 ($p = .003$). Messzeitpunkt 2 und Messzeitpunkt 3 unterscheiden sich nicht signifikant voneinander ($p = 1.000$).

7.2.4. Ergebnisse der Gruppenvergleiche im residualen Lernzuwachs (H6)

Reflexionskompetenztest – Deklaratives und konditionales Wissen

Im Folgenden werden die Ergebnisse des residualen Lernzuwachses des Reflexionskompetenztests der drei Gruppen präsentiert (Abbildung 70, Tabelle 146):

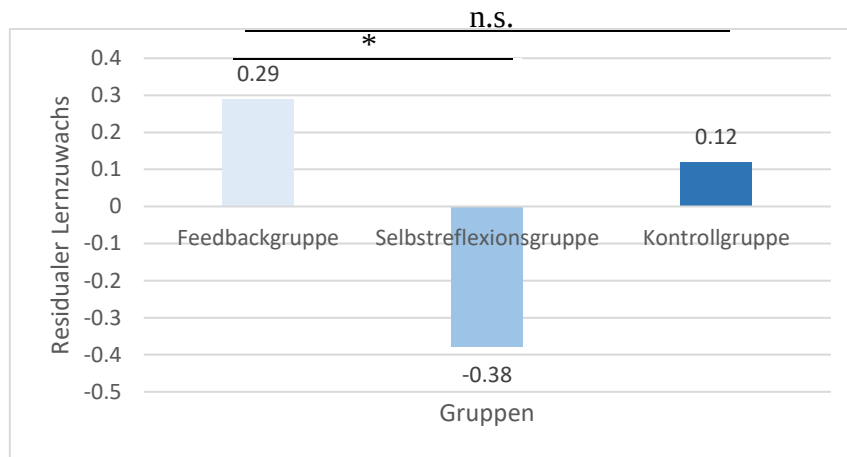


Abbildung 70: Hauptstudie – Gruppenvergleiche – Reflexionskompetenztest – Gesamtstichprobe

Tabelle 146: Hauptstudie – Gruppenvergleiche – Reflexionskompetenztest – Deskriptive Statistik – Gesamtstichprobe

Residualer Lernzuwachs	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	.29	0.76	1.67	-1.22
Selbstreflexionsgruppe	21	-.38	0.99	1.32	-2.04
Kontrollgruppe	17	.12	1.13	1.32	-2.63

Um zu überprüfen, ob die Feedbackgruppe (IG1) der Selbstreflexionsgruppe (IG2) und der Kontrollgruppe (KG) überlegen ist, folgt ein Vergleich der standardisierten Residuen mittels *t*-Tests für unabhängige Stichproben. Da letztere normalverteilt vorliegen, wird auf einen nichtparametrischen Test verzichtet.

IG1 vs. IG2: $t(35) = 2.43$; $p = .020$; $d = -0.8$

IG1 vs. KG: $t(35) = 0.54$; $p = .591$; n.s.; $d = -0.2$

Kodiermanual – Prozedurales Wissen

Im Folgenden werden die Ergebnisse des residualen Lernzuwachses im prozeduralen Wissen vorgestellt (Abbildung 71):

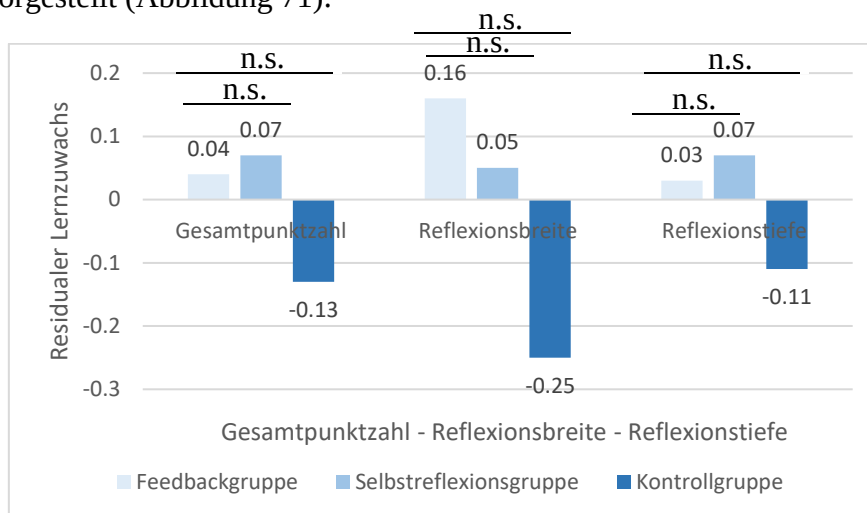


Abbildung 71: Hauptstudie – Gruppenvergleiche residualer Lernzuwachs – Reflexionskompetenz – Kodiermanual – Gesamtpunktzahl, Reflexionsbreite, Reflexionstiefe

Gesamtpunktzahl (Tabelle 147)

Tabelle 147: Hauptstudie – Gruppenvergleiche – Deskriptive Statistik – Kodiermanual – Gesamtpunktzahl

Gesamtpunktzahl	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	.04	1.02	3.50	-1.21
Selbstreflexionsgruppe	21	.07	1.01	2.49	-1.61
Kontrollgruppe	17	-.13	0.98	2.32	-1.40

Reflexionsbreite (Tabelle 148)

Tabelle 148: Hauptstudie – Gruppenvergleiche – Deskriptive Statistik – Kodiermanual – Reflexionsbreite

Reflexionsbreite	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	.16	0.80	1.62	-1.31
Selbstreflexionsgruppe	21	.05	1.03	1.37	-2.06
Kontrollgruppe	17	-.25	1.15	1.62	-3.21

Reflexionstiefe (Tabelle 149)

Tabelle 149: Hauptstudie – Gruppenvergleiche – Deskriptive Statistik – Kodiermanual – Reflexionstiefe

Reflexionstiefe	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	.03	1.03	3.63	-1.20
Selbstreflexionsgruppe	21	.07	1.00	2.45	-1.58

Kontrollgruppe	17	-.11	.98	2.38	-1.27

Da die standardisierten Residuen nicht alle normalverteilt vorliegen, werden je nach Ergebnis nichtparametrische und parametrische Tests berechnet. Zur Ergänzung der nichtparametrischen Tests sind die Ergebnisse der parametrischen Berechnungen im Anhang zu finden.

Gesamtpunktzahl

Nichtparametrischer Test (Mann-Whitney-U-Test)

IG1 vs. IG2: $z = -0.24$; $p = .814$; n.s.; $r = 0.0$

IG1 vs. KG: $z = 0.47$; $p = .478$; n.s.; $r = 0.1$

Parametrischer Test (Tabelle 296, Tabelle 297)

Reflexionsbreite

Parametrischer Test (t -Test für unabhängige Stichproben)

IG1 vs. IG2: $t(39) = 0.37$; $p = .716$; n.s.; $d = -0.1$

IG1 vs. KG: $t(35) = 1.26$; $p = .216$; n.s.; $d = -0.4$

Reflexionstiefe

Nichtparametrischer Test (Mann-Whitney-U-Test)

IG1 vs. IG2: $z = -0.26$; $p = .794$; n.s.; $r = 0.0$

IG1 vs. KG: $z = -0.73$; $p = .478$; n.s.; $r = 0.1$

Parametrischer Test (Tabelle 302, Tabelle 303)

Obwohl deskriptiv die Kontrollgruppe in allen Bereichen des prozeduralen Wissens hinter den Interventionsgruppen zurückbleibt, ist der Effekt nicht statistisch signifikant. Ebenso wenig lassen sich statistisch relevante Unterschiede zwischen den beiden Interventionsgruppen ausmachen.

Experimentell-fachdidaktisches Wissen – Gruppenvergleiche im residualen Lernzuwachs

Im Folgenden wird der residualer Lernzuwachs im fachdidaktischen Wissen vorgestellt (Abbildung 72, Tabelle 150):

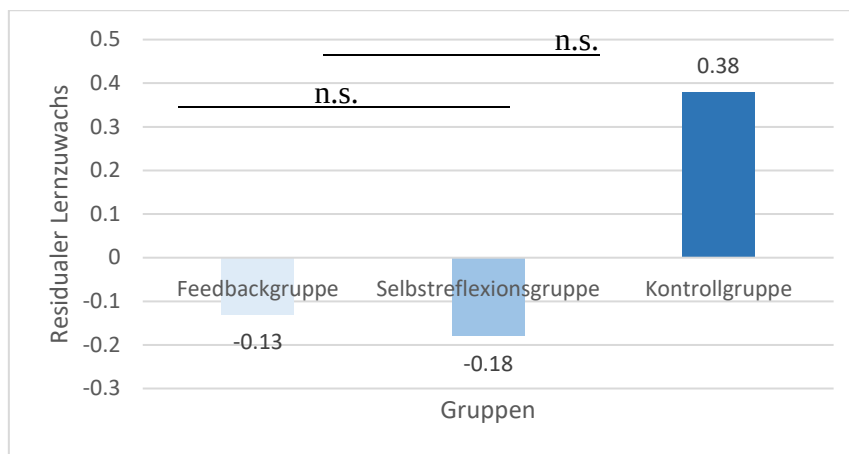


Abbildung 72: Hauptstudie – Gruppenvergleiche residualer Lernzuwachs – fachdidaktisches Wissen

Tabelle 150: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche residualer Lernzuwachs – fachdidaktisches Wissen

Residualer Lernzuwachs Fachdidaktisches Wissen	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	-.13	1.07	1.98	-2.10
Selbstreflexionsgruppe	21	-.18	1.10	1.55	-2.51
Kontrollgruppe	17	.38	0.65	1.77	-.59

Um zu überprüfen, ob es statistisch signifikante Unterschiede in den Gruppen gibt, werden die standardisierten Residuen mittels Mann-Whitney-U Test verglichen.

Nichtparametrischer Test (Mann-Whitney-U-Test)

IG1 vs. IG2: $z = -0.25$; $p = .804$; n.s.; $r = 0.0$

IG1 vs. KG: $z = 1.30$; $p = .198$; n.s.; $r = 0.2$

Parametrischer Test (Tabelle 307, Tabelle 308)

Die beiden Interventionsgruppen bleiben im fachdidaktischen Wissen deskriptiv betrachtet hinter den Erwartungen zurück, wohingegen die Kontrollgruppe die Erwartungen übertrifft. Statistisch sind jedoch keine Unterschiede zu erkennen.

Fachwissen – Gruppenvergleiche im residualen Lernzuwachs

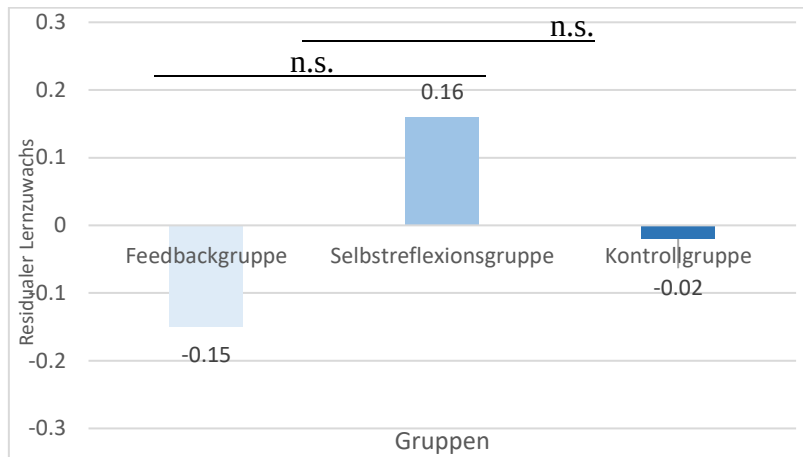


Abbildung 73: Hauptstudie – Gruppenvergleiche residualer Lernzuwachs – Fachwissen

Tabelle 151: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche residualer Lernzuwachs – Fachwissen

Standardisierte Residuen					
CK-Test	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	-.15	.92	1.26	-1.81
Selbstreflexionsgruppe	21	.16	1.05	2.11	-1.76
Kontrollgruppe	17	-.02	1.02	1.74	-1.51

Um zu überprüfen, ob es statistisch signifikante Unterschiede zwischen den Gruppen gibt, werden die standardisierten Residuen mittels Mann-Whitney-U Test verglichen.

Nichtparametrischer Test (Mann-Whitney-U-Test)

IG1 vs. IG2: $z = -0.93$; $p = .354$; n.s.; $r = 0.1$

IG1 vs. KG: $z = -0.46$; $p = .662$; n.s.; $r = 0.1$

Parametrischer Test (Tabelle 312, Tabelle 313)

Beim Fachwissenstest kann kein statistisch signifikanter Unterschied zwischen den Gruppen im residualen Lernzuwachs festgestellt werden. Deskriptiv betrachtet ist hier die IG2, also die Selbstreflexionsgruppe, besser als die zu erwartenden Ergebnisse. Im Gegensatz hierzu fällt die Feedbackgruppe ab.

Selbstwirksamkeit – Gruppenvergleiche der residualen Veränderung

Selbstwirksamkeit – Gesamtpunktzahl (Abbildung 74, Tabelle 152)

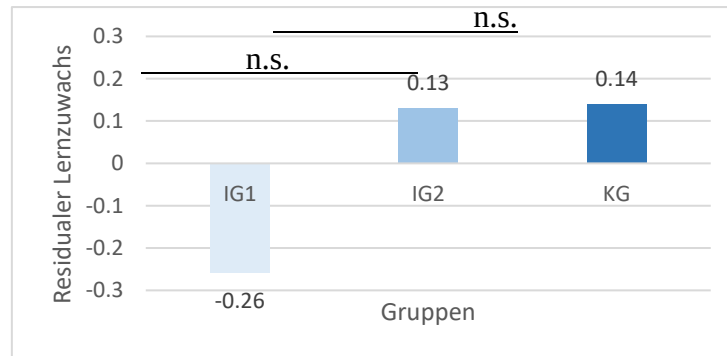


Abbildung 74: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Gesamtpunktzahl

Tabelle 152: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Gesamtpunktzahl

Selbstwirksamkeit Gesamt	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	-.26	1.11	1.71	-2.81
Selbstreflexionsgruppe	21	.13	0.98	1.67	-1.66
Kontrollgruppe	17	.14	0.85	1.44	-1.14

Da die standardisierten Residuen alle normalverteilt vorliegen, werden parametrische Verfahren berechnet.

Parametrischer Test (*t*-Test für unabhängige Stichproben)

IG1 vs. IG2: $t(39) = -1,22$; $p = .229$; n.s.; $d = 0.4$

IG1 vs. KG: $t(35) = -1,24$; $p = .222$; n.s.; $d = 0.4$

Bei der Gesamtpunktzahl der Selbstwirksamkeitserwartung sind keine signifikanten Unterschiede in den Gruppen zu erkennen.

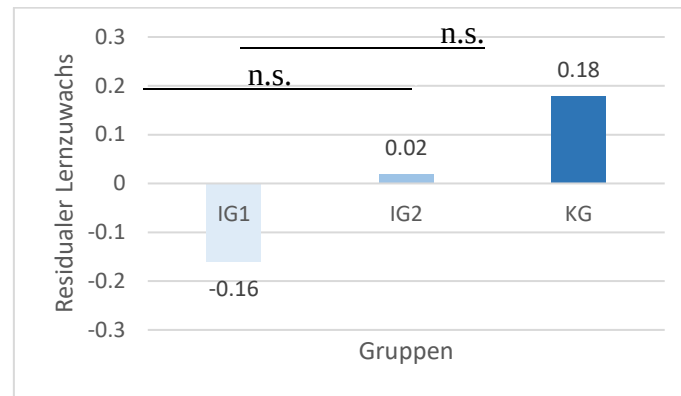
Selbstwirksamkeit – Planung von Experimenten (Abbildung 75, Tabelle 153)

Abbildung 75: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Planung von Experimenten

Tabelle 153: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Planung von Experimenten

SR Planung von Experimenten	N	M	SD	Max.	Min.
Feedbackgruppe	20	-.16	1.16	3.19	-1.90
Selbstreflexionsgruppe	21	.02	0.78	1.81	-.94
Kontrollgruppe	16	.18	1.05	2.48	-1.08

Parametrischer Test (t -Test für unabhängige Stichproben)

IG1 vs. IG2: $t(39) = -0.57$; $p = .576$; n.s.; $d = 0.2$

IG1 vs. KG: $t(34) = -0.90$; $p = .374$; n.s.; $d = 0.3$

Es ist kein signifikanter Unterschied der Gruppen im Konstrukt Selbstwirksamkeit – Planung von Experimenten zu erkennen.

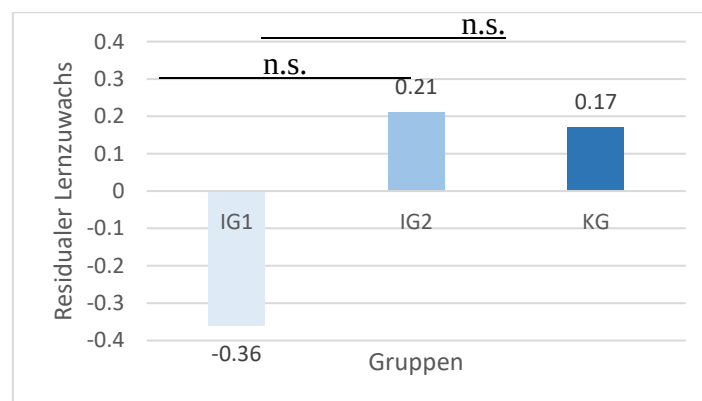
Selbstwirksamkeit – Durchführung von Experimenten

Abbildung 76: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Durchführung von Experimenten

Tabelle 154: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Durchführung von Experimenten

Durchführung von Experimenten	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	-.36	1.07	1.14	-2.04
Selbstreflexionsgruppe	21	.21	1.04	1.63	-2.01
Kontrollgruppe	16	.17	0.72	1.12	-1.12

Parametrischer Test (*t*-Test für unabhängige Stichproben)

IG1 vs. IG2: $t(39) = -1,72$; $p = .093$; n.s.; $d = 0.5$

IG1 vs. KG: $t(33) = -1,77$; $p = .085$; n.s.; $d = 0.6$

Es sind keine signifikanten Unterschiede im Konstrukt Selbstwirksamkeit – Durchführung von Experimenten zu erkennen.

Selbstwirksamkeit – Didaktische Rekonstruktion unterrichtsrelevanter Inhalte (Abbildung 77/Tabelle 155)

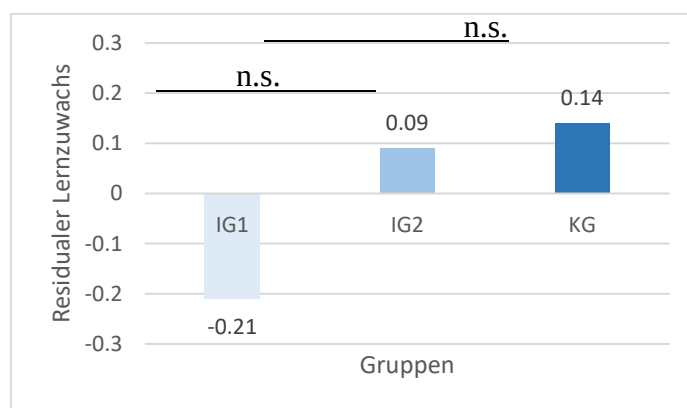


Abbildung 77: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Didaktische Rekonstruktion

Tabelle 155: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Didaktische Rekonstruktion

Didaktische Rekonstruktion	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	-.21	0.90	1.43	-1.83
Selbstreflexionsgruppe	21	.09	0.95	2.27	-1.60
Kontrollgruppe	17	.14	1.15	3.04	-1.55

Parametrischer Test (*t*-Test für unabhängige Stichproben)

IG1 vs. IG2: $t(39) = -1,07$; $p = .292$; n.s.; $d = 0.3$

IG1 vs. KG: $t(35) = -1,04$; $p = .304$; n.s.; $d = 0.3$

Es sind keine signifikanten Unterschiede im Konstrukt Selbstwirksamkeit – Didaktische Rekonstruktion unterrichtsrelevanter Inhalte zwischen den Gruppen zu erkennen.

Selbstwirksamkeit – Umsetzung der didaktischen Rekonstruktion im Unterricht (Abbildung 78, Tabelle 156)



Abbildung 78: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Umsetzung der didaktischen Rekonstruktion

Tabelle 156: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Umsetzung der didaktischen Rekonstruktion

Umsetzung der didaktischen Rekonstruktion	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	-.09	1.14	2.39	-2.07
Selbstreflexionsgruppe	21	-.07	0.99	1.56	-1.79
Kontrollgruppe	17	.19	0.83	1.56	-1.79

Parametrischer Test (*t*-Test für unabhängige Stichproben)

IG1 vs. IG2: $t(39) = -0,05$; $p = .963$; n.s.; $d = 0.0$

IG1 vs. KG: $t(35) = -0,82$; $p = .419$; n.s.; $d = 0.3$

Es sind keine signifikanten Unterschiede Im Konstrukt Selbstwirksamkeit – Umsetzung der didaktischen Rekonstruktion im Unterricht zu erkennen.

Unterrichtswirksamkeit

Zur Berechnung der Wirksamkeit des gehaltenen Unterrichts werden die standardisierten Residuen der Wissenstestergebnisse der Schülerinnen und Schüler betrachtet. Hierbei ist nur ein Vergleich aus IG1 und IG2 möglich, da die Kontrollgruppe die Unterrichtsstunden lediglich theoretisch plante und ihre Stunden nicht vor Schülerinnen und Schülern unterrichtete (Abbildung 79, Tabelle 157).

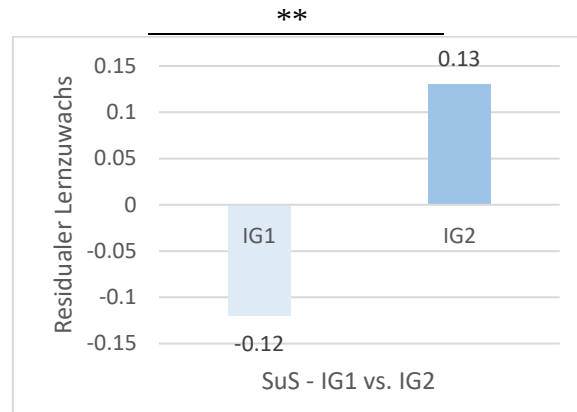


Abbildung 79: Hauptstudie – Gruppenvergleiche – Residualer Lernzuwachs der Unterrichtswirksamkeit SuS – Gesamtstichprobe

Tabelle 157: Hauptstudie – Deskriptive Statistik – Gruppenvergleich – Residualer Lernzuwachs der Unterrichtswirksamkeit SuS – Gesamtstichprobe

SR CK SuS	N	M	SD	Max.	Min.
Feedbackgruppe	255	-.12	0.96	2.67	-2.26
Selbstreflexionsgruppe	245	.13	1.02	2.82	-2.67

Da der residuale Lernzuwachs der Schülerinnen und Schüler normalverteilt vorliegt, werden parametrische Tests berechnet.

Parametrischer Test (*t*-Test für unabhängige Stichproben)

IG1 vs. IG2: $t(498) = -2,85$; $p = .004$; $d = 0.3$

Vergleich der Gruppen über die drei Messzeitpunkte (Abbildung 80, Tabelle 158):

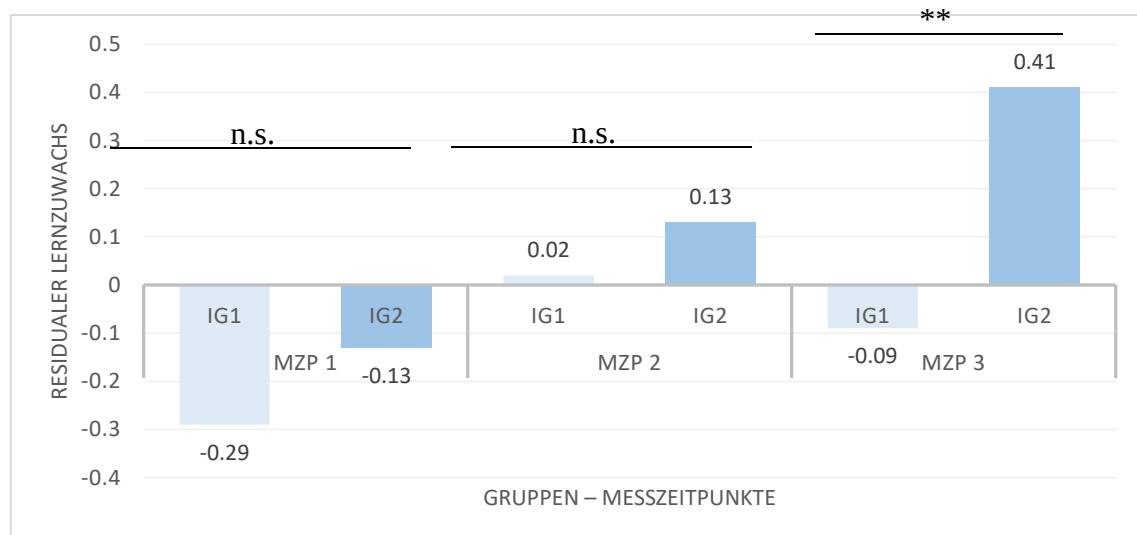


Abbildung 80: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Unterrichtswirksamkeit – Vergleich der Gruppen über die drei Messzeitpunkte

Tabelle 158: Hauptstudie – Deskriptive Statistik – Residualer Lernzuwachs der Unterrichtswirksamkeit – Vergleich der Gruppen über die drei Messzeitpunkte

		<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
MZIP 1	Feedbackgruppe	90	-.29	0.98	2.67	-2.22
	Selbstreflexionsgruppe	80	-.13	0.94	1.90	-2.67
MZIP 2	Feedbackgruppe	81	.02	0.93	2.12	-2.07
	Selbstreflexionsgruppe	90	.13	0.97	2.82	-1.96
MZIP 3	Feedbackgruppe	84	-.09	0.96	2.21	-2.26
	Selbstreflexionsgruppe	75	.41	1.11	2.82	-1.84

Parametrischer Test (t-Test für unabhängige Stichproben)

MZIP1: $t(168) = -1,06$; $p = .292$; n.s.; $d = 0.2$

MZIP2: $t(169) = -0,74$; $p = .463$; n.s.; $d = 0.1$

MZIP3: $t(157) = -3,04$; $p = .003$; $d = 0.5$

Die Gruppen unterscheiden sich im dritten Messzeitpunkt signifikant voneinander, wobei die Selbstreflexionsgruppe eine höhere Unterrichtswirksamkeit bezüglich des residualen Lernzuwachses der Schülerinnen und Schüler im Fachwissen aufweist.

Entwicklung der Unterrichtswirksamkeit der zwei Interventionsgruppen:

Feedbackgruppe (Abbildung 81, Tabelle 159):

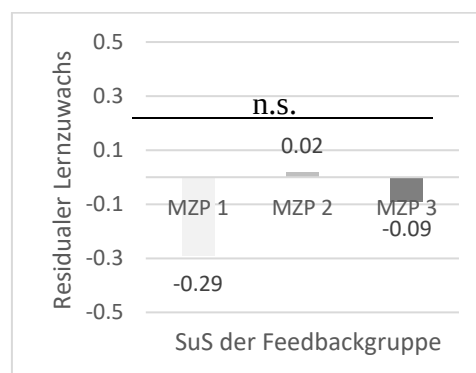


Abbildung 81: Hauptstudie – Feedbackgruppe – Residuale Entwicklung der Unterrichtswirksamkeit über drei Messzeitpunkte

Tabelle 159: Hauptstudie – Feedbackgruppe – Deskriptive Statistik – Residuale Entwicklung der Unterrichtswirksamkeit über drei Messzeitpunkte

Feedbackgruppe	N	M	SD	Max.	Min.
MZP 1	90	-.29	0.98	2.67	-2.22
MZP 2	81	.02	0.93	2.12	-2.07
MZP 3	84	-.09	0.96	2.21	-2.26

Die univariate Varianzanalyse lässt in der Feedbackgruppe keine signifikanten Unterschiede in den drei Messzeitpunkten erkennen ($F(2,254) = 2,31$; $p = .102$; n.s.; $\eta^2 = .02$)

Selbstreflexionsgruppe (Abbildung 82, Tabelle 160):

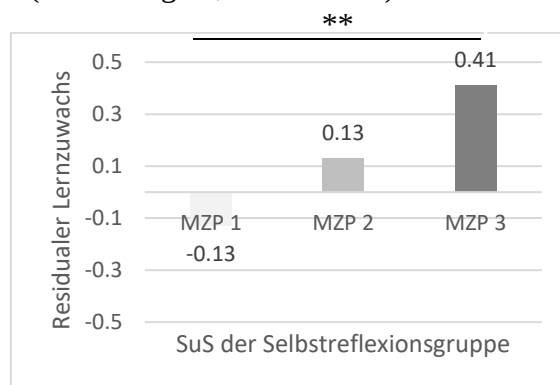


Abbildung 82: Hauptstudie – Selbstreflexionsgruppe – Residuale Entwicklung der Unterrichtswirksamkeit über drei Messzeitpunkte

Tabelle 160: Hauptstudie – Selbstreflexionsgruppe – Deskriptive Statistik – Residuale Entwicklung der Unterrichtswirksamkeit über drei Messzeitpunkte

Selbstreflexionsgruppe	N	M	SD	Max.	Min.
MZP 1	80	-.13	0.94	1.90	-2.67
MZP 2	90	.13	0.97	2.82	-1.96
MZP 3	75	.41	1.11	2.82	-1.84

Die univariate Varianzanalyse der Selbstreflexionsgruppe ist über die drei Messzeitpunkte signifikant ($F(2,244) = 5,69$; $p = .004$; $\eta^2 = 0.05$). Der Post-Hoc-Test zeigt lediglich zwischen erstem und drittem Messzeitpunkt signifikante Unterschiede ($p = .003$).

7.3. Weitere Ergebnisse

7.3.1. Experimentell-fachdidaktisches Wissen – Vergleich IG2 vs. KG

Experimentell-fachdidaktisches Wissen – Vergleich von Selbstreflexionsgruppe und Kontrollgruppe im residualen Lernzuwachs

In Kapitel 7.2 wurde bereits gezeigt, dass zwischen Feedbackgruppe und der Selbstreflexionsgruppe, ebenso wie zwischen der Feedbackgruppe und der Kontrollgruppe keine signifikanten Unterschiede zu erkennen sind. Die gerichtete Hypothese 6 konnte somit nicht bestätigt werden. Da jedoch der Unterschied zwischen Selbstreflexionsgruppe und Kontrollgruppe größer ausfiel als zwischen Feedbackgruppe und Kontrollgruppe werden ergänzende Rechnungen vorgenommen: Um zu überprüfen, ob es signifikante Unterschiede zwischen Selbstreflexionsgruppe und Kontrollgruppe gibt, werden die standardisierten Residuen mittels t -Test für unabhängige Stichproben verglichen. Da letztere normalverteilt vorliegen, kann auf einen nichtparametrischen Test verzichtet werden.

IG2 vs. KG: $t(36) = -1,46$; $p = .152$; n.s.; $d = 0.5$

Es lassen sich auch zwischen IG2 und KG keine Unterschiede erkennen.

7.3.2. Korrelation Reflexionskompetenztest mit fachdidaktischem Wissen und Fachwissen unter Ausschluss der IG1 – Vergleich IG2 vs. KG

Da der Reflexionskompetenztest sich als ungeeignet für die Feedbackgruppe herausgestellt hat, wurden sämtliche Korrelationen erneut nur mit Vergleich Selbstreflexionsgruppe und Kontrollgruppe gerechnet. Im Folgenden werden die Ergebnisse der Korrelationen des residualen Lernzuwachses im fachdidaktischen Wissen und Fachwissen berichtet: Da die residualen Lernzuwächse nicht immer normalverteilt vorliegen, werden die Korrelationen nach Pearson und/oder die Korrelationen nach Spearman berichtet (Tabelle 161).

Tabelle 161: Hauptstudie – Korrelationen des residualen Lernzuwachses – fachdidaktisches Wissen und Fachwissen mit Kodiermanual – Ausschluss IG1

Residualer Lernzuwachs	Fachdidaktisches Wissen	Fachwissen
Reflexionskompetenztest	$r(38) = .33$; $p = .047$ $r_s(38) = .41$; $p = .010$;	$r(38) = .24$; $p = .139$; n.s. $r_s(38) = .30$; $p = .070$; n.s.

Bei Ausschluss der Feedbackgruppe korreliert der residuale Lernzuwachs des Reflexionskompetenztests mit dem residualen Lernzuwachs des fachdidaktischen Wissens. Weitere Korrelationen können nicht festgestellt werden.

7.3.3. Kodiermanual – Perspektivenübernahme und Literaturbezug

Ergebnisse zur Perspektivenübernahme – Gesamt:

Es folgen zunächst die Ergebnisse der Gesamtheit der möglichen Perspektivenübernahmen (Tabelle 162).

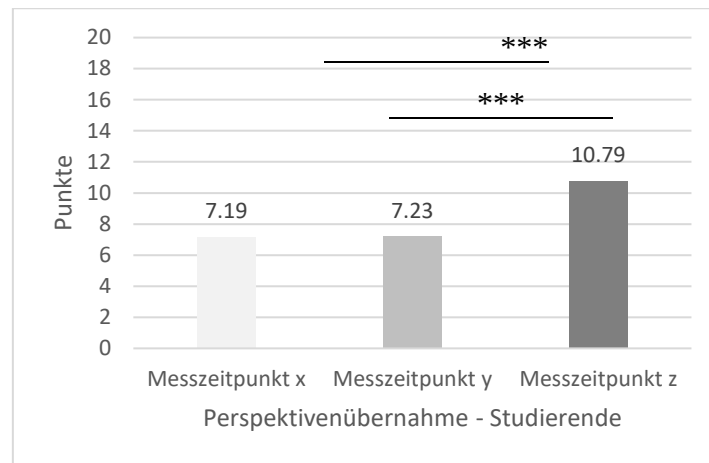


Abbildung 83: Hauptstudie – Kodiermanual – Mittlere Scores – Perspektivenübernahme – Gesamtstichprobe – Messzeitpunkte x,y,z

Tabelle 162: Hauptstudie – Kodiermanual – Deskriptive Statistik – Perspektivenübernahme – Gesamtstichprobe – Messzeitpunkte x,y,z

Gesamtstichprobe	N	M	SD	Max.	Min.
Messzeitpunkt x	57	7.19	5.43	32.00	1.00
Messzeitpunkt y	57	7.23	4.30	18.00	1.00
Messzeitpunkt z	58	10.79	5.12	25.00	1.00

Da die Perspektivenübernahme nicht normalverteilt vorliegt, werden nichtparametrische Tests berechnet.

Wissenszuwachs:

Nichtparametrischer Test (Wilcoxon-Test):

$$Z_{MZPx-MZPy} = -0.15; p = .881; \text{n.s.}; r = 0.01;$$

$$Z_{MZPx-MZPz} = -4.17; p < .001; r = 0.4;$$

$$Z_{MZPy-MZPz} = -5.18; p < .001; r = 0.5;$$

Parametrischer Test (Tabelle 339):

Bei der Perspektivenübernahme ist ein hochsignifikanter Wissenszuwachs von MZPx auf MZPz und von MZPy auf MZPz zu beobachten.

Ergebnisse zur Perspektivenübernahme – Aufteilung in Gruppen:

Im Folgenden werden die Ergebnisse der Gesamtheit der Perspektivenübernahme unter Aufteilung der Gruppen berichtet (Abbildung 84, Tabelle 163, Tabelle 164, Tabelle 165).

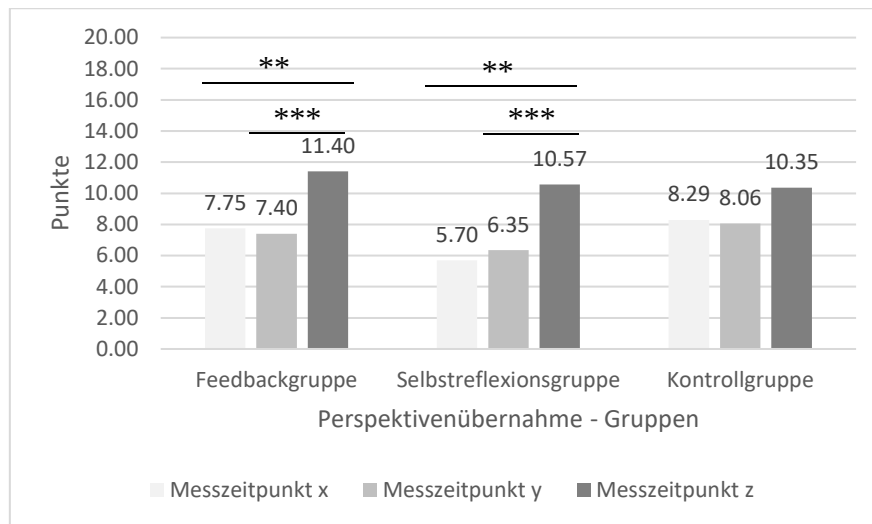


Abbildung 84: Hauptstudie – Kodiermanual – Mittlere Scores – Perspektivenübernahme – Gruppenaufteilung – Messzeitpunkte x,y,z

Tabelle 163: Hauptstudie – Kodiermanual – Deskriptive Statistik – Perspektivenübernahme – Feedbackgruppe – Messzeitpunkte x,y,z

Feedbackgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	20	7.75	6.70	32.00	1.00
Messzeitpunkt y	20	7.40	4.36	18.00	2.00
Messzeitpunkt z	20	11.40	3.91	19.00	6.00

Tabelle 164: Hauptstudie – Kodiermanual – Deskriptive Statistik – Perspektivenübernahme – Selbstreflexionsgruppe – Messzeitpunkte x,y,z

Selbstreflexionsgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	20	5.70	4.17	17.00	1.00
Messzeitpunkt y	20	6.35	4.55	16.00	1.00
Messzeitpunkt z	21	10.57	6.06	25.00	1.00

Tabelle 165: Hauptstudie – Kodiermanual – Deskriptive Statistik – Perspektivenübernahme – Kontrollgruppe – Messzeitpunkte x,y,z

Kontrollgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	17	8.29	4.96	17.00	1.00
Messzeitpunkt y	17	8.06	3.98	16.00	3.00
Messzeitpunkt z	17	10.35	5.33	20.00	2.00

Die Kontrollgruppe als diejenige Gruppe, welche von Beginn an Fremdmaterial beurteilte, sich also in fremde Personen hineinversetzen musste, erreicht im MZPx und MZPy den höchsten Mittelwert. Im MZPz können die Interventionsgruppen den Unterschied aufholen. Der Lernzuwachs ist in diesen beiden Gruppen hoch signifikant.

Wissenszuwachs:

Nichtparametrischer Test (Wilcoxon-Test) (Tabelle 166):

Tabelle 166: Hauptstudie – Kodiermanual – Perspektivenübernahme – Wissenszuwachs – nichtparametrisch – Gruppenaufteilung – Messzeitpunkte x,y,z

Gruppe	Messzeitpunkte	<i>z</i>	<i>p</i>	<i>r</i>
Feedbackgruppe <i>n</i> = 20	MZPx-MZPy	-0.33	= .738; n.s.	0.1
	MZPx-MZPz	-2.86	= .004	0.5
	MZPy-MZPz	-3.23	=.001	0.5
Selbstreflexionsgruppe <i>n</i> = 21	MZPx-MZPy	-0.69	= .492; n.s.	0.1
	MZPx-MZPz	-2.82	= .005	0.4
	MZPy-MZPz	-3.62	< .001	0.6
Kontrollgruppe <i>n</i> = 17	MZPx-MZPy	-0.40	= .690; n.s.	0.1
	MZPx-MZPz	-1.59	= .111; n.s.	0.3
	MZPy-MZPz	-1.77	= .077; n.s.	0.3

Parametrischer Test (Tabelle 341, Tabelle 343, Tabelle 345):

Residualer Lernzuwachs – Perspektivenübernahme

Im Folgenden werden die Ergebnisse des residualen Lernzuwachses der Perspektivenübernahme vorgestellt (Tabelle 167):

Tabelle 167: Hauptstudie – Kodiermanual – Perspektivenübernahme – residualer Lernzuwachs – Gruppenaufteilung

Residualer Lernzuwachs Perspektivenübernahme	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	0.08	0.66	1.57	-0.92
Selbstreflexionsgruppe	20	0.08	1.25	3.08	-1.80
Kontrollgruppe	17	-0.18	1.01	1.91	-1.49

Parametrischer Test (*t*-Test für unabhängige Stichproben)

IG1 vs. IG2: $t(38) = -0.00$; $p = .998$; n.s.; $d = 0.0$

IG1 vs. KG: $t(35) = 0.024$; $p = .362$; n.s.; $d = -0.3$

Es finden sich bzgl. der Perspektivenübernahme keine signifikanten Unterschiede zwischen Feedbackgruppe und Selbstreflexion- bzw. Kontrollgruppe.

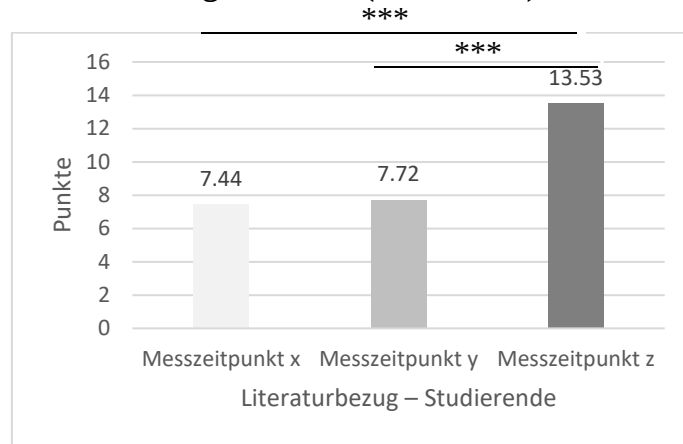
Ergebnisse zum Literaturbezug – Gesamt (Tabelle 168):

Abbildung 85: Hauptstudie – Kodiermanual – Literaturbezug – Mittlere Scores – Gesamtstichprobe – Messzeitpunkte x,y,z

Tabelle 168: Hauptstudie – Kodiermanual – Literaturbezug – Deskriptive Statistik – Gesamtstichprobe – Messzeitpunkte x,y,z

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Messzeitpunkt x	57	7.44	5.74	34.00	1.00
Messzeitpunkt y	57	7.72	4.92	21.00	1.00
Messzeitpunkt z	58	13.53	6.95	32.00	1.00

Wissenszuwachs

Da die drei Messzeitpunkte nicht normalverteilt vorliegen, werden nichtparametrische berechnet.

Nichtparametrischer Test (Wilcoxon-Test):

$$Z_{\text{MZPx-MZPy}} = -0.38; p = .707; \text{n.s.}; r = 0.04;$$

$$Z_{\text{MZPx-MZPz}} = -4.97; p < .001; r = 0.5;$$

$$Z_{\text{MZPy-MZPz}} = -5.75; p < .001; r = 0.5;$$

Parametrischer Test (Tabelle 350):

Beim Konstrukt Literaturbezug ist ebenfalls ein hochsignifikanter Wissenszuwachs von MZPx auf MZPz und von MZPy auf MZPz zu beobachten.

Ergebnisse zum Literaturbezug – Aufteilung in Gruppen (Abbildung 86, Tabelle 169, Tabelle 170, Tabelle 171):

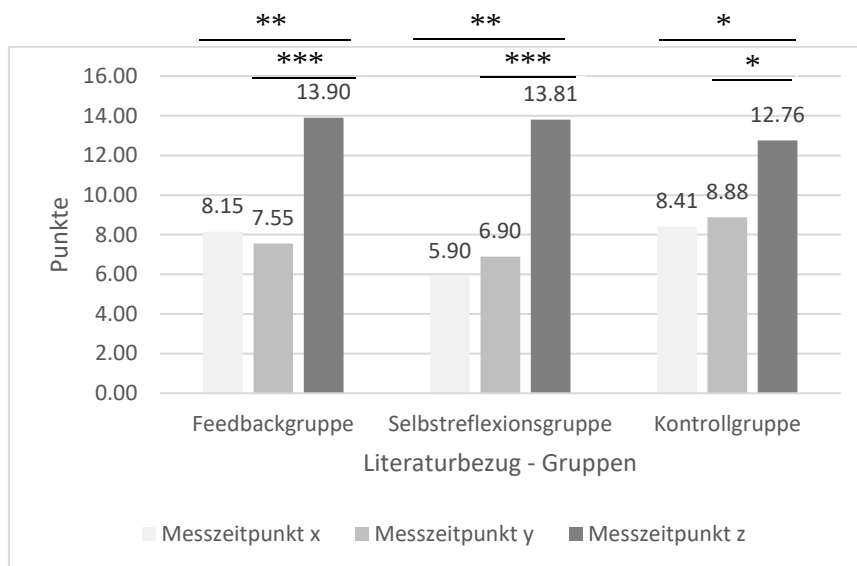


Abbildung 86: Hauptstudie – Kodiermanual – Literaturbezug – Mittlere Scores – Gruppenaufteilung – Messzeitpunkte x,y,z

Tabelle 169: Hauptstudie – Kodiermanual – Literaturbezug – Deskriptive Statistik – Feedbackgruppe – Messzeitpunkte x,y,z

Feedbackgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	20	8.15	7.20	34.00	1.00
Messzeitpunkt y	20	7.55	4.77	21.00	2.00
Messzeitpunkt z	20	13.90	5.10	26.00	6.00

Tabelle 170: Hauptstudie – Kodiermanual – Literaturbezug – Deskriptive Statistik – Selbstreflexionsgruppe – Messzeitpunkte x,y,z

Selbstreflexionsgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	20	5.90	4.40	17.00	1.00
Messzeitpunkt y	20	6.90	5.24	19.00	1.00
Messzeitpunkt z	21	13.81	7.99	27.00	1.00

Tabelle 171: Hauptstudie – Kodiermanual – Literaturbezug – Deskriptive Statistik – Kontrollgruppe – Messzeitpunkte x,y,z

Kontrollgruppe	N	M	SD	Max.	Min.
Messzeitpunkt x	17	8.41	5.11	18.00	1.00
Messzeitpunkt y	17	8.88	4.78	18.00	3.00
Messzeitpunkt z	17	12.76	7.77	32.00	2.00

Wissenszuwachs:

Nichtparametrischer Test (Wilcoxon-Test) (Tabelle 172):

Tabelle 172: Hauptstudie – Kodiermanual – Literaturbezug – Wissenszuwachs – nicht-parametrisch – Gruppenaufteilung – Messzeitpunkte x,y,z

Gruppe	Messzeitpunkte	<i>z</i>	<i>p</i>	<i>r</i>
Feedbackgruppe <i>n</i> = 20	MZPx-MZPy	-0.57	= .569; n.s.	0.1
	MZPx-MZPz	-3.14	= .002	0.5
	MZPy-MZPz	-3.78	< .001	0.6
Selbstreflexionsgruppe <i>n</i> = 21	MZPx-MZPy	-0.68	= .499; n.s.	0.1
	MZPx-MZPz	-3.16	= .002	0.5
	MZPy-MZPz	-3.70	< .001	0.6
Kontrollgruppe <i>n</i> = 17	MZPx-MZPy	-0.86	= .864; n.s.	0.1
	MZPx-MZPz	-2.28	= .023	0.4
	MZPy-MZPz	-2.23	= .026	0.4

Parametrischer Test (Tabelle 354, Tabelle 355, Tabelle 356):

Feedbackgruppe und Selbstreflexionsgruppe lernen von MZPy auf MZPz hoch signifikant dazu.

Residualer Lernzuwachs – Literaturbezug (Tabelle 173)

Tabelle 173: Hauptstudie – Gruppenvergleiche – Deskriptive Statistik – Kodiermanual – Literaturbezug

Literaturbezug	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Max.</i>	<i>Min.</i>
Feedbackgruppe	20	0.01	0.70	1.70	-1.0
Selbstreflexionsgruppe	20	0.13	1.18	2.09	-1.68
Kontrollgruppe	17	-0.17	1.07	2.73	-1.39

Nichtparametrischer Test (Mann-Whitney-U-Test):

IG1 vs. IG2: $z = -0.15$; $p = .883$; n.s.; $r = 0.0$

IG1 vs. KG: $z = -1.17$; $p = .244$; n.s.; $r = 0.2$

Parametrischer Test (Tabelle 354)

7.3.4. Extremgruppenvergleiche

Um Extremgruppenvergleiche ziehen zu können, wurde das fachdidaktische Wissen-Pre-Testergebnis mithilfe des Medians dichotomisiert. Anschließend wurde der Einfluss auf die Steigerung der Reflexionskompetenz durch eine Anova berechnet. Ebenso wurde anhand des Fachwissen-Pre-Testergebnisses dichotomisiert und den Einfluss auf die Steigerung der Reflexionskompetenz untersucht. Die Steigerung der Reflexionskompetenz wurde einmal über den Wissenstest und einmal über das Kodiermanual im Rahmen einer Anova mit MZP 3 bzw. MZPz als abhängige Variable und MZP 1 bzw. MZPx als unabhängiger Variable erhoben. Faktor war die nach PCK bzw. CK dichotomisierte Stichprobe.

Einfluss des Fachdidaktischen-Tests auf die Reflexionskompetenz:

Reflexionskompetenztest: $F(2,57) = 0,01; p = .947; \text{n.s.}; \eta^2 = 0.00$

Gesamtpunktzahl Kodiermanual: $F(2,57) = 0,47; p = .496; \text{n.s.}; \eta^2 = 0.01$

Einfluss des Fachwissen-Tests auf die Reflexionskompetenz:

Reflexionskompetenztest: $F(2,57) = 1,08; p = .303; \text{n.s.}; \eta^2 = 0.01$

Gesamtpunktzahl Kodiermanual: $F(2,57) = 0,11; p = .744; \text{n.s.}; \eta^2 = 0.00$

Weder das fachdidaktische Wissen-Pre-Testergebnis noch das Fachwissen-Pre-Testergebnis haben Einfluss auf die Steigerung der Reflexionskompetenz. Somit kann davon ausgegangen werden, dass sich das konzipierte Seminar sowohl für leistungsstarke als auch leistungsschwächere Studierende eignet.

Um das Ergebnis abzusichern, wurde eine weitere Anova gerechnet: Als unabhängige Variablen dienten der jeweilige Reflexionskompetenzwert zu MZP 1 bzw. MZPx, der fachdidaktische Wissen-Pre-Wert oder der Fachwissen-Pre-Wert. Als abhängige Variable wurden der jeweilige Reflexionskompetenzwert zu MZP 2 bzw. MZPz verwendet.

Fachdidaktisches Wissen-Pre-Test:

Reflexionskompetenztest: $F(2,57) = 0,02; p = .884; \text{n.s.}; \eta^2 = 0.00$

Gesamtpunktzahl Kodiermanual: $F(2,57) = 0,00; p = .978; \text{n.s.}; \eta^2 = 0.00$

Fachwissen-Pre-Test:

Reflexionskompetenztest: $F(2,57) = 0,02; p = .891; \text{n.s.}; \eta^2 = 0.00$

Gesamtpunktzahl Kodiermanual: $F(2,57) = 1.03; p = .315; \text{n.s.}; \eta^2 = 0.02$

Demnach lässt sich folgern, dass die Reflexionskompetenz unabhängig vom Vorwissen im Bereich fachdidaktisches Wissen und Fachwissen ist.

8. Diskussion, Reflexion und Ausblick

In Kapitel 8.1. werden die Ergebnisse der Hauptstudie hinsichtlich der Forschungsfrage 2 auf Basis der Hypothesen überprüft und kritisch reflektiert. In Kapitel 8.2 folgen die Hypothesen zur Forschungsfrage 2. Ergänzend dazu werden weitere Ergebnisse, wie etwa Extremgruppenberechnungen, diskutiert und ein kurzer Ausblick auf mögliche Folgestudien gegeben.

8.1. Gestaltung eines universitären Seminars zur Förderung der Reflexionskompetenz von Chemielehramtsstudierenden (Forschungsfrage 1)

Im Folgenden werden die Hypothesen mittels der in Kapitel 7 dargestellten Ergebnisse auf ihre Gültigkeit getestet:

H1 Durch das universitäre Seminar wird die Reflexionskompetenz der Chemielehramtsstudierenden gesteigert.

Reflexionskompetenztest (konditionales und deklaratives Wissen)

In der Gesamtstichprobe werden im Post-Test höchst signifikant bessere Werte erzielt als im Pre-Test. Ebenso kann die Nachhaltigkeit bestätigt werden, da auch die Ergebnisse des Follow-Up-Tests höchst signifikant über den Pre-Test Ergebnissen liegen. Vergleicht man die einzelnen Gruppen (Feedbackgruppe, Selbstreflexions- und Kontrollgruppe), so werden höchst signifikante Ergebnisse in allen drei Gruppen erreicht. Sowohl von Pre-Test auf Post-Test als auch von Pre-Test auf Follow-Up-Test. Somit kann H1 hinsichtlich des konditionalen und deklarativen Wissens bestätigt werden. Eine Erklärung für diese einheitliche Verbesserung könnte die gemeinsame Inputphase zu Beginn des Semesters sein, in welcher die Studierenden Grundkenntnisse zum Thema Reflexion erlernen. Die hohen Effektstärken lassen sich womöglich auf die dreimaligen schriftlichen Reflexionen zurückführen, in welchen die Studierenden ihre Handlungen in (selbst) gehaltenen Unterrichtsstunden einschätzen müssen und so auch Kenntnisse im Bereich des konditionalen Wissens verbessern.

Kodiermanual (prozedurales Wissen)

Ebenso kann im prozeduralen Wissen – in der Gesamtstichprobe – eine höchst signifikante Steigerung in der Gesamtpunktzahl von Messzeitpunkt x auf Messzeitpunkt z festgestellt werden. Zwischen Messzeitpunkt x und Messzeitpunkt y wird kein Wissenszuwachs gemessen. Der Gruppenvergleich zeigt lediglich bei der Feedbackgruppe und der Selbstreflexionsgruppe eine signifikante Steigerung von Messzeitpunkt x (Beginn des Seminars) auf Messzeitpunkt z (am Ende des Seminars). Die Kontrollgruppe erreicht keine signifikante Steigerung. Dies könnte daran liegen, dass die Kontrollgruppe über die Unterrichtsstunden der Feedbackgruppe und damit fremde Stunden schriftlich reflektiert. Möglicherweise sind Reflexionen in Bezug auf das für Unterrichtsplanungen erforderliche Wissen weniger wirksam, wenn sie nicht mit eigenen Unterrichtserfahrungen verbunden werden (Nitsche, 2014).

Es wurde auf eine unterschiedliche Gewichtung der verschiedenen Bereiche der Reflexionstiefe wie in anderen Studien verzichtet (Kempin et al., 2019) und diese mit jeweils einem Punkt bewertet. Aufgrund der wiederholten Reflexion über die thematisch gleiche, dreimalig verbesserte Unterrichtsstunde treten die in anderen Studien zum Teil höher bewerteten Facetten, wie beispielsweise Konsequenzen und Verbesserungsvorschläge, im

Verlauf des Semesters weniger auf (Abbildung 60). Dies lässt sich durch die stetig positivere Bewertung der gehaltenen Stunden erklären. Die negativen Aspekte nehmen entsprechend ab. So werden folglich auch weniger Konsequenzen aus Fehlverhalten gezogen bzw. Verbesserungsvorschläge gemacht. Somit ist eine unterschiedliche Gewichtung diverser Facetten in diesem Studiendesign nicht sinnvoll.

Unterteilt man die Gesamtpunktzahl in Reflexionsbreite und Reflexionstiefe, können hinsichtlich der Reflexionsbreite weder in der Gesamtstichprobe noch bei Gruppenaufteilung signifikante Ergebnisse erreicht werden. Möglicherweise stellen die ca. sechs bis sieben reflektierten Kategorien (von maximal 10 möglichen) eine leistbare und sinnvolle Breite dar, die mit dem primären Anspruch, die Qualität und Tiefe der Reflexionen zu verbessern, noch vereinbar ist. In der Reflexionstiefe erreichen die Studierenden sowohl von Messzeitpunkt y (Mitte des Seminars) auf Messzeitpunkt z (Ende des Seminars) als auch von Messzeitpunkt x (Beginn des Seminars) auf Messzeitpunkt z höchst signifikante Steigerungen. Der Gruppenvergleich fällt auch hier zugunsten der Feedbackgruppe und der Selbstreflexionsgruppe aus. Diese können beide von Messzeitpunkt x auf Messzeitpunkt z signifikante Steigerungen verzeichnen. Die Feedbackgruppe erreicht von Messzeitpunkt y auf Messzeitpunkt z höchst signifikante Werte, was aber auf eine Verschlechterung von Messzeitpunkt x auf Messzeitpunkt y zurückzuführen ist. Die Kontrollgruppe verbessert sich zwar rein deskriptiv, es können aber keine signifikanten Unterschiede gefunden werden. H1 wird ebenso in Bezug auf das prozedurale Wissen bestätigt.

H2 Die Reflexionskompetenz der Studierenden lässt sich mithilfe des entwickelten Kodiermanuals objektiv, reliabel und valide erfassen.

Die sozialwissenschaftliche Methodenlehre teilt die Gütekriterien der Inhaltsanalyse in Maße der Reliabilität und der Validität ein. Durch die Berechnungen der Inter-coder-reliabilität wird zusätzlich die Objektivität bestimmt (Mayring, 2010). In dieser Studie können alle Kriterien für die Qualität der Inhaltsanalyse nach Krippendorff (1980) nachgewiesen werden. Diese beinhaltet unter anderem die Validität im engeren Sinne, zu welcher materialorientiert die semantische Gültigkeit und die Stichprobengültigkeit zählen. Ergebnisorientiert wurden korrelative Gültigkeiten und Vorhersagegültigkeit getestet, prozessorientiert wurde die Konstruktgültigkeit geprüft. Die Reliabilität ist unterteilt in Stabilität, Reproduzierbarkeit und Exaktheit. Alle drei Punkte konnten erfolgreich getestet werden. Die guten Werte beruhen womöglich auf der Empfehlung von Kember et al. (1999), welche sich für projektinterne Kodierer einsetzen. Dies ist vor allem an Kodierer 3 zu erkennen, welcher erst nach intensiver Besprechung vergleichbare Werte wie Kodierer 1 und 2, die bereits früh in das Projekt einbezogen worden sind, liefert. Mit überschaubarem Schulungsaufwand scheint das Manual geeignet zu sein, die Reflexionskompetenz von Studierenden zu erfassen. H2 kann bestätigt werden.

8.2. Zusammenhänge zwischen der Förderung der Reflexionskompetenz und diversen didaktischen Aspekten (Forschungsfrage 2)

H3a Mittels des universitären Seminars lassen sich das experimentell-fachdidaktische Wissen und das Fachwissen zum Themenfeld Stofftrennung und -identifikation der Studierenden steigern.

Die Gesamtstichprobe kann in dieser Studie keine signifikante Steigerung im Bereich des experimentell-fachdidaktischen Wissens (PCK) verzeichnen. Im Fachwissen zum Themenfeld Stofftrennung und -identifikation können jedoch sowohl von Pre-Test auf Post-Test als auch von Pre-Test auf Follow-Up-Test hochsignifikante Ergebnisse erzielt werden. Damit kann im Bereich des Fachwissens ein nachhaltiger Lernerfolg nachgewiesen werden. Somit kann das abgeänderte Studiendesign im Vergleich zur Studie von Anthofer (2017) im Bereich fachdidaktisches Wissen nicht an die Ergebnisse anknüpfen. Der Bereich Fachwissen wird jedoch bestätigt.

Der mangelnde Lernzuwachs im experimentell-fachdidaktischen Wissen könnte primär darin begründet liegen, dass der Fokus des Seminars auf der Steigerung der Reflexionskompetenz als einem speziellen Teil des fachdidaktischen Wissens von Lehrkräften gelegen hat, sodass eine Steigerung im experimentell-fachdidaktischen Wissen nicht im Maß wie bei (Anthofer, 2017) zu erwarten war. Ein weiterer Grund kann darin liegen, dass sich Studierende stärker auf die Klasse als auf einzelne Schülerinnen und Schüler fokussieren, während sie insgesamt schlechte Kompetenzen im Monitoring aufweisen (Barth, 2017; Emmer & Gerwels, 2006; Wolff, van den Bogert, Jarodzka & Boshuizen, 2015). Studierende in den ersten selbst erteilten Unterrichtsstunden sind durch die Konzentration auf den Einsatz der Medien (Powerpoint, Dokumentenkamera, Experimente, ...) stark beansprucht, dass parallel ablaufende Prozesse nicht umfänglich wahrgenommen werden (Denner & Gesenhues, 2013). So zeigen sie eine zu späte oder gar fehlende Wahrnehmung von Störungen im Unterricht, was eine weitreichende, tiefe Reflexion und somit eine Verbesserung im Bereich fachdidaktisches Wissen erschweren könnte (Housner & Griffey, 1985). Werden Probleme wahrgenommen, fokussieren sie eher statisch-oberflächlich. Im Vergleich hierzu finden geübte Lehrerinnen und Lehrer dynamische Tiefenstrukturen (Berliner, 1988; Bromme, 2014; Carter, Cushing, Sabers, Stein & Berliner, 1988). Ebenso leidet nach Emmer und Gerwels (2006) die Beurteilungskompetenz. Studierende ohne Unterrichtsexpertise schaffen es nicht, die Komplexität im Klassenzimmer zu überblicken. So werden Unterrichtsaspekte isoliert betrachtet, Zusammenhänge übersehen (Carter et al., 1988; Harrington et al., 1996; Wolff et al., 2015) und oftmals lediglich verkürzte Problemreflexionen durchgeführt (Swanson, O'Connor & Cooney, 1990). Santagata und Guarino (2011) konnten zudem zeigen, dass auch das fehlende Wissen über Lehr-Lern-Strategien zu Schwierigkeiten beim Überlegen von Handlungsalternativen führt. Es werden daher nur wenige Alternativen für Unterrichtsstrategien entwickelt (Barth, 2017; Housner & Griffey, 1985). Studierende bevorzugen schnelle Lösungen (Swanson et al., 1990). Diese fehlenden Tiefenstrukturen beim Wahrnehmen von fachdidaktischen Problemen und Generieren von Handlungsalternativen könnten der Grund dafür sein, dass sich die Studierenden vermehrt auf den Bereich des Fachwissens fokussierten. Hypothese H3a kann somit nur für das Fachwissen bestätigt werden.

H3b Der Lernzuwachs bezüglich der Reflexionskompetenz korreliert positiv mit dem experimentell-fachdidaktischen Wissen und dem Fachwissen zum Themenfeld Stofftrennung und -identifikation.

Für die Gesamtstichprobe gilt, dass weder der Reflexionskompetenztest noch das Kodiermanual mit dem experimentell-fachdidaktischen Wissen oder dem Fachwissen zum Themenfeld Stofftrennung und -identifikation korrelieren. Aufgrund eines fehlenden Lernzuwachses im experimentell-fachdidaktischen Wissen ist eine fehlende Korrelation zwi-

schen diesem Konstrukt und der gesteigerten Reflexionskompetenz in der Gesamtstichprobe plausibel. Die nicht gefundenen Korrelationen können möglicherweise auch auf die Passung des Reflexionskompetenztests für die drei Vergleichsgruppen zurückgeführt werden. Die Reliabilität des Reflexionskompetenztests ist unter Ausschluss der Feedbackgruppe wesentlich besser, jedoch wurden die Berechnungen aufgrund einer besseren Vergleichbarkeit unter Einbezug aller Gruppen durchgeführt. Bei Ausschluss der Feedbackgruppe kann die Korrelation des Reflexionskompetenztests mit dem experimentell-fachdidaktischen Wissen nachgewiesen werden. Erklärungen könnten sein, dass das ePCK durch den Wissensstand der Lehrperson, durch die Wahl der Unterrichtsmethoden und -strategien, in der Begründung für spezifische pädagogische Handlungen und in der Integration multipler Faktoren pädagogischen Handelns (z. B. Kenntnisse der Schülerinnen und Schüler oder Bewertungskompetenz) sichtbar wird (Carlson & Daehler, 2019). Genau diese Begründungen werden von den Studierenden in den schriftlichen Reflexionen verlangt.

Im Fachwissen war keine Korrelation mit der Reflexionskompetenz erkennbar. Grund könnte sein, dass in diesem Kodiermanual die Facetten der Reflexionskompetenz unabhängig von der fachlichen bzw. der fachdidaktischen Richtigkeit bepunktet wurden. Somit konnten möglicherweise fachlich falsche Reflexionen im Bereich Reflexionskompetenz trotzdem Punkte erzielen. Auch wenn sich ein Zusammenhang zwischen fachdidaktischem Wissen bzw. Fachwissen und Reflexionskompetenz theoretisch begründen lässt, könnte das für eine sinnvolle Reflexion der Unterrichtsplanungen und -durchführungen für den Anfangsunterricht erforderliche Wissen so gering sein, dass nahezu jeder Studierende dieses beherrscht. Aus den beiden genannten Gründen ist eine Korrelation der beiden Bereiche weniger wahrscheinlich.

H3b muss folglich gesondert betrachtet werden. Eine Korrelation der Reflexionskompetenz mit dem fachdidaktischen Wissen ist z. T. nachweisbar. Für das Fachwissen ist H3b zu falsifizieren.

H4 Der Lernzuwachs bezüglich der Reflexionskompetenz der Studierenden korreliert positiv mit ...

H4.1 ... ihren Zielorientierungen im Studium.

H4.2 ... ihren fachspezifischen Überzeugungen.

H4.3 ... ihren Selbstwirksamkeitserwartungen.

Die Zielorientierungen bzw. die Bestrebungen im Studium weisen innerhalb der Gesamtstichprobe keine signifikanten Veränderungen auf. Die fachspezifischen Überzeugungen, welche sich in die drei Unterpunkte Chemie als Wissenschaft, Unterrichten von Chemie in der Schule und Lernen von Chemie in der Schule unterteilen, können hingegen über die drei Messzeitpunkte hochsignifikante Verbesserungen verzeichnen. Lediglich der Punkt Lernen von Chemie in der Schule relativiert sich zum Follow-Up-Test. Insgesamt gesehen haben sich die Einstellungen zur Chemie als Wissenschaft und zum Unterrichten von Chemie in der Schule nachhaltig über die Gesamtstichprobe verbessert. Die Selbstwirksamkeitserwartungen können sich in allen Bereichen – Planung von Experimenten, Durchführung von Experimenten, didaktische Rekonstruktion und Umsetzung der didak-

tischen Rekonstruktion – vom Pre-Test zum Post-Test signifikant bis hochsignifikant verbessern. Alle Subkonstrukte weisen vom Pre-Test zum Follow-Up-Test eine hochsignifikante Steigerung auf. Die Einstellungsänderungen scheinen damit nachhaltig zu sein.

Korrelationen der Zielorientierungen mit den Wissensbereichen der Reflexionskompetenz:

Bei der Berechnung von Korrelationen weisen die Zielorientierungen lediglich mit der Gesamtpunktzahl des Kodiermanuals eine sehr schwache Korrelation auf und dies auch nur bei den Berechnungen für nicht normalverteilte Korrelationen. Deswegen wird dieses Ergebnis nicht weiter interpretiert.

Korrelationen der fachspezifischen Überzeugungen mit den Wissensbereichen der Reflexionskompetenz:

Der Unterpunkt Chemie als Wissenschaft korreliert stark mit der Reflexionsbreite des Kodiermanuals. Der Punkt Unterrichten von Chemie korreliert sowohl mit der Reflexionstiefe als auch mit der Gesamtpunktzahl des Kodiermanuals. Der letzte Unterpunkt Lernen von Chemie korreliert negativ mit dem Reflexionskompetenztest.

Zur möglichen Erklärung: Carlson und Daehler (2019) stellen fest, dass ein videobasierter Reflexions- und Feedbackzirkel motivierende Auswirkungen auf Zielorientierungen und fachspezifische Überzeugungen mit sich bringt. Gründe, dass die Zielorientierungen nicht mit den Bereichen der Reflexionskompetenz korrelieren, könnten sein, dass die Items der Zielorientierungen teilweise zu weit von der Unterrichtssituation und der damit verbundenen Reflexion entfernt waren. Beispielimts:

In meinem Studium strebe ich danach, ...

... meinen Kommilitonen zu zeigen, dass ich mehr weiß als andere.

... meinem Dozenten zu beweisen, dass ich mehr weiß als andere.

... vor meinem Dozenten zu verbergen, wenn ich etwas weniger gut kann als andere.

... dass ich mir selbst nicht vorhalten muss, ich würde mein Studium weniger gut bewältigen als andere.

... mir selbst nicht eingestehen zu müssen, wenn ich etwas weniger gut kann als andere.

Das in der vorliegenden Arbeit gewählte Studiendesign zielt nicht auf diese Facetten der Zielorientierung ab. Folglich kann weder eine Verbesserung noch eine eindeutige Korrelation festgestellt werden.

Bei Betrachtung der Items zum Punkt fachspezifische Überzeugungen fällt auf, dass die Unterpunkte unterschiedlich, jedoch in der Gesamtheit mit allen Facetten der Reflexionskompetenz korrelieren. Insbesondere ist die Korrelation des Unterpunkts Unterrichten von Chemie mit der Gesamtpunktzahl bzw. der Reflexionstiefe interessant. Dies könnte auf die inhaltliche Übereinstimmung zurückzuführen sein. Als Beispiel werden drei Items aufgeführt:

„Der Chemieunterricht sollte sich an Alltagsphänomenen orientieren.“

„Eine Lehrkraft sollte Dinge ausprobieren, da einem der Unterricht dann selbst mehr Freude bereitet.“

„SuS sollen im Chemieunterricht experimentelles Arbeiten lernen.“

Dies sind Aspekte, die zahlreich in den schriftlichen Reflexionen erwähnt wurden, sodass eine starke Korrelation auf diese Weise plausibel ist.

Der Unterpunkt Lernen von Chemie in der Schule korreliert negativ mit dem Reflexionskompetenztest. Möglicherweise kann dies mit der Cognitive Load Theory von Sweller (1994) erklärt werden. Der Inhalt der Items fragt stark offene Lernsituationen ab und somit für Lehranfänger Situationen, in denen sie gefühlt die Kontrolle im Unterricht verlieren können. Der Cognitive Overload durch die Fülle von Aufgaben im Unterricht könnte evtl. dazu führen, dass unsichere Studierende ein eher angeleitetes Unterrichtsmuster als angenehmer empfinden als offene Unterrichtssituationen, in denen Schülerinnen und Schüler selbstverantwortlich lernen bzw. experimentieren. Im Wissenstest zur Reflexionskompetenz können sie jedoch das Theoriewissen anwenden, da in den Item-Beispielen fremde Lehrpersonen Ziel der Reflexionen.

Korrelationen der Selbstwirksamkeitserwartungen mit den Wissensbereichen der Reflexionskompetenz:

Lediglich der Unterpunkt Umsetzung der didaktischen Rekonstruktion korreliert negativ mit dem Reflexionskompetenztest. Eine mögliche Erklärung liefern Denner und Gesenhues (2013), welche herausfanden, dass Studierende Sicherheit durch den Verweis auf Regeln und Prozeduren erlangen, dies aber zu einer Abnahme der Verantwortungsübernahme für das eigene Handeln zur Folge hat. Somit wäre es denkbar, dass sich einerseits ein Teil der Studierenden zwar im Reflexionskompetenztest immer mehr auf Regeln und Theoriewissen stützen konnte, dies aber bei Items wie beispielsweise

„Auch wenn ich meinen Chemieunterricht spontan an Interessen der Schülerinnen und Schüler anpasse, gelingt es mir immer, eine sinnvolle Reihenfolge beizubehalten.“

oder

„Im Chemieunterricht bin ich stets in der Lage, eine Erklärung weiter zu vereinfachen, die die Schülerinnen und Schüler nicht verstehen, auch wenn ich darauf nicht gefasst war.“

einen Verlust an Selbstwirksamkeitserwartungen durch negative Erfahrungswerte in den drei gehaltenen Stunden zur Folge gehabt haben könnte. Das eigene Handeln wird evtl. nach den gehaltenen Stunden kritischer betrachtet worden sein. Als Folge könnte dies zu Lasten der Verantwortungsübernahme für einen reibungslosen Unterrichtsverlauf gegangen sein. Betrachtet man die andere Seite, wäre es möglich, dass diejenigen Studierenden ihre Selbstwirksamkeitserwartungen steigerten, welche bereits erlernte Handlungsmuster weiter vertieften und weiteres Theoriewissen ausblendeten, also auch keine neuen Methoden ausprobierten. So könnte ein Anstieg der Selbstwirksamkeitserwartungen, während eines Rückgangs an Theoriewissen erklärt werden. Folglich kann die Studie von Dohrmann und Nordmeier (2018) für dieses Studiendesign bestätigt werden, welche besagt, dass Unterrichten in Lehr-Lern-Laboren mit anschließender Reflexion zu einer Steigerung der Selbstwirksamkeitserwartung führt. Korrelationen mit der Reflexionskompetenz finden jedoch nur zum Teil Bestätigung. H4 kann also nur für die Unterhypothese H4.2 bestätigt werden.

H5 Der Lernzuwachs im Fachwissen der jeweils an den Messzeitpunkten unterrichteten Schülerinnen und Schüler steigert sich von Messzeitpunkt 1 über Messzeitpunkt 2 bis zu Messzeitpunkt 3.

Für die Gesamtstichprobe kann nachgewiesen werden, dass die Schülerinnen und Schüler von Pre-Test auf Post-Test und auch von Pre-Test auf Follow-Up-Test einen hochsignifikanten Lernzuwachs verzeichnen. Die Effektstärken bleiben annähernd auf dem gleichen Niveau. Im residualen Lernzuwachs ergibt sich eine Steigerung vom ersten Messzeitpunkt über den zweiten Messzeitpunkt hin zum dritten Messzeitpunkt. Somit werden die Ergebnisse der Studie von Anthofer (2017) bestätigt. H5 kann für diese Studie als gültig erachtet werden.

H6 Die Leistungen (der residuale Lernzuwachs) der Feedbackgruppe übertreffen (übertrifft) die (den), der Selbstreflexionsgruppe und der Kontrollgruppe in den Bereichen ...

H 6.1 Reflexionskompetenz

H 6.2 Lernzuwachs im experimentell-fachdidaktischen Wissen

H 6.3 Lernzuwachs im Fachwissen

H 6.4 Selbstwirksamkeit

H 6.5 Unterrichtswirksamkeit

H6.1

Reflexionskompetenz – Reflexionskompetenztest (deklaratives und konditionales Wissen)

Im Reflexionskompetenztest kann die Feedbackgruppe einen signifikant höheren Lernzuwachs als die Selbstreflexionsgruppe verzeichnen. Jedoch finden sich keine Unterschiede zur Kontrollgruppe. Gründe hierfür könnten in der individuellen Förderung der Feedbackgruppe liegen. Nach Carlson und Daehler (2019) lässt sich das pPCK durch professionelles Feedback fördern. Dies beinhaltet das persönliche Wissen und die eigene Expertise einer Lehrerin/eines Lehrers über das Unterrichten eines bestimmten Fachgebiets, welches sich aus Erfahrungen ergibt, was wiederum direkte Auswirkungen auf des ePCK hat, das unter anderem die Reflexion beinhaltet. Somit konnte womöglich das persönliche Wissen der kumulierten Erfahrungen aus den Unterrichtsstunden in der Feedbackgruppe eher beispielsweise auf das konditionale Wissen des Reflexionskompetenztests übertragen werden. Ein möglicher Grund für das annähernd gleich gute Ergebnis von Feedback- und Kontrollgruppe könnte sein, dass diese durch die nicht selbst zu haltenden Unterrichtsstunden kognitiv weniger überlastet waren als die unterrichtende Gruppe und so eine nicht erfolgte Förderung durch fehlendes Feedback kompensieren konnte. Eine mögliche Überforderung durch die Fülle der Aufgaben einer Lehrperson könnte für die Selbstreflexionsgruppe nachteilig gewesen sein (Sweller, 1994).

Reflexionskompetenz – Kodiermanual (prozedurales Wissen)

Es zeigen sich keine signifikanten Unterschiede im residualen Lernzuwachs zwischen Selbstreflexionsgruppe, Feedbackgruppe und Kontrollgruppe, weder in der Gesamtpunktzahl noch in den Teilbereichen Reflexionstiefe und Reflexionsbreite. Es scheint also, dass das prozedurale Wissen im Themenfeld Reflexionskompetenz unabhängig vom Feedback bzw. unabhängig von der Unterrichtspraxis ist. Damit ist das Ergebnis auf den ersten Blick widersprüchlich zum Befund, dass sich Feedback- und Selbstreflexionsgruppe bzgl. der Reflexionstiefe verbessern, während sich die Kontrollgruppe nicht signifikant steigern kann. Die nicht belegbaren Unterschiede im residualen Lernzuwachs deuten darauf hin, dass die Effekte nicht so groß sind, um zweifelsfrei nachgewiesen zu werden. Eine Diskussion über mögliche Ursachen erscheint zu diesem Zeitpunkt verfrüht.

H 6.1 kann nur zum Teil bestätigt werden. Die Feedbackgruppe überzeugt lediglich im Reflexionskompetenztest (deklaratives und konditionales Wissen), zusammen mit der Kontrollgruppe.

H6.2

Lernzuwachs im experimentell-fachdidaktischen Wissen

Im experimentell-fachdidaktischen Wissen können keine signifikanten Unterschiede zwischen den Gruppen verzeichnet werden, ebenso wenig ein signifikanter Anstieg des Wissens innerhalb der Gesamtstichprobe. Wie bereits erwähnt, sind die Studierenden in den ersten Unterrichtsstunden mit sehr vielen verschiedenen Aspekten des Unterrichtens beschäftigt, sodass eine Fokussierung auf experimentell-fachdidaktische Aspekte, als Teil des fachdidaktischen Wissens, nicht möglich scheint (Housner & Griffey, 1985; Swanson et al., 1990). Carlson und Daehler (2019) berichten von einer möglichen Weiterentwicklung des pPCKs durch eine Vertiefung bzw. eine Reflexion der Selbstwirksamkeitserwartungen, Zielorientierungen und fachspezifischen Überzeugungen. Dies kann durch Feedback unterstützt werden. In dieser Studie lag der Fokus des Feedbacks primär auf der Weiterentwicklung der Unterrichtseinheit und weniger auf der des experimentell-fachdidaktischen Wissens. H6.2 kann nicht bestätigt werden.

H6.3

Lernzuwachs im Fachwissen

Im Fachwissen kann kein signifikanter Unterschied zwischen den Gruppen gezeigt werden. Die fehlenden Unterschiede in den Interventionsgruppen könnten womöglich auf die interne und externe Feedbackstruktur dieser Studie zurückzuführen sein. Es werden beide Interventionsgruppen durch die Unterrichtsstunden mit internem Feedback konfrontiert. Nach Copeland et al. (1993) könnte durch Lücken im Fachwissen, welche beispielsweise durch Schülernachfragen identifiziert werden, eine Problemidentifikation stattfinden. In beiden Interventionsgruppen ist es möglich, durch die drei Messzeitpunkte Lösungen zu generieren, also die fehlenden Informationen zum nächsten Messzeitpunkt nachzuarbeiten und somit die dritte Oberstufe des Modells, Lösung ausprobieren, durchzuführen. Die Feedbackgruppe erhält zwar zusätzlich externes Feedback und somit die Möglichkeit etwaige fachliche Fehler in der Durchführung der Stunden aufdecken zu können, welche in der Selbstreflexionsgruppe womöglich unbemerkt bleiben. Die Lerngelegenheiten sind aber möglicherweise aufgrund der im Seminar ohnehin vorgesehenen Planungsstunden

für alle Gruppen in gewisser Weise ähnlich. Ebenso ist es möglich, dass die Kontrollgruppe durch das Beobachten der fremden Lehrpersonen Vorteile in der Analyse problematischer Ereignisse hat und somit kritische Situationen in Zusammenhang mit fachlichen Problemen auch ohne internes und externes Feedback identifizieren kann (Seidel, Stürmer, Blomberg, Kobarg & Schwindt, 2011). Dies bestätigen auch Kleinknecht und Poschinski (2014), welche die Effektivität der Reflexionen eigener und fremder Unterrichtsvideos untersuchten. In ihren Ergebnissen berichten sie, dass die emotionale Distanz bei Analyse von Fremdvideos eher zu einer vertieften Reflexion kritischer Situationen führen kann, als die Reflexion eigener Unterrichtsvideos. H6.3 kann nicht bestätigt werden.

H6.4

Selbstwirksamkeit

Weder in der Gesamtpunktzahl noch in den Einzelfacetten, Planung von Experimenten, Durchführung von Experimenten, didaktische Rekonstruktion unterrichtsrelevanter Inhalte und Umsetzung der didaktischen Rekonstruktion im Unterricht, können signifikante Unterschiede festgestellt werden. Gründe für dieses Ergebnis könnten im Studiendesign zu finden sein. Öhlschlegel-Haubrock et al. (2014) fanden heraus, dass bei einem Vergleich der Wirksamkeit eines Mentoringkonzepts mit einem dozentenorientierten Lehrformats das Mentorenformat eine höhere Einschätzung der Selbstwirksamkeitserwartung beobachtet ließ. Zentraler Faktor dieser Studie war die Förderung des selbstgesteuerten Lernens durch eine adäquate Lernumgebung, welche den Studierenden ermöglichte, selbst aktiv zu werden und kooperativ zu lernen.

Auch wenn die Feedbackgruppe der vorliegenden Arbeit durch den Experten spezielle Unterstützung hat und die Kontrollgruppe keinen Unterricht halten darf, werden alle Gruppen auf Basis der Learning Study (Nilsson, 2014) unterstützt, selbstgesteuert und aktiv zu lernen. So erhalten alle die gleiche Inputphase, alle Probanden haben die gleiche Zeit für ihre geplanten Unterrichtsstunden zu recherchieren und sich vorzubereiten. Es sind immer drei Messzeitpunkte einer Unterrichtsstunde zum selben Thema und diese Stunden werden immer gleich reflektiert und im Anschluss überarbeitet. Evtl. ist es möglich, dass durch die Förderung des selbstgesteuerten Lernens die Auswirkungen des Feedbacks nivelliert. Dies würde die Studie von Dohrmann und Nordmeier (2018) bestätigen, welche die Auswirkungen des Unterrichtens in Lehr-Lern-Laboren auf Reflexion und Selbstwirksamkeitserwartungen untersuchten und auch hier positive Ergebnisse in Bezug auf die Selbstwirksamkeit erzielten. Gold et al. (2017) führten ihre positiven Ergebnisse zu den Selbstwirksamkeitserwartungen auf die Analyse von fremden und eigenen Unterrichtsvideos zurück, was in dieser Studie auch auf alle Probanden zutrifft. Es könnte also möglicherweise sein, dass der videobasierte Aufbau des Seminars die Selbstwirksamkeitserwartungen aller drei Gruppen gleich beeinflusst. H6.4 kann nicht bestätigt werden.

H6.5

Unterrichtswirksamkeit

Die Schülerinnen und Schüler der Selbstreflexionsgruppe übertrifft die der Feedbackgruppe hinsichtlich des Lernzuwachses signifikant. Werden die einzelnen Messzeitpunkte verglichen, sind keine signifikanten Unterschiede in den beiden Gruppen im ersten

und zweiten Messzeitpunkt zu verzeichnen. Im dritten Messzeitpunkt erreichen die Schülerinnen und Schüler der Selbstreflexionsgruppe deutlich bessere Ergebnisse als die Schülerinnen und Schüler der Feedbackgruppe. Betrachtet man die einzelnen Gruppen, ist bei der Feedbackgruppe keine signifikante Entwicklung zu beobachten. Die Selbstreflexionsgruppe hingegen steigert sich vom ersten über den zweiten hin zum dritten Messzeitpunkt signifikant. Möglicherweise ist der Grund für dieses Ergebnis in einer Kombination aus Studiendesign, dem Cognitive Overload sowie externem und internem Feedback zu finden. Die Probanden aus beiden Interventionsstudien planen ihre Unterrichtsstunden zum ersten Messzeitpunkt mit vergleichbaren Bedingungen. Dies ist im ersten Messzeitpunkt zu sehen, welcher keine Unterschiede in den Gruppen aufweist. Mehrere Studien bestätigen, dass Studierenden in ihren ersten Unterrichtsstunden mit der Fülle an Unterrichtsprozessen überfordert sind und auf oberflächliche Probleme fokussieren (Barth, 2017; Denner & Gesenhues, 2013; Dreyfus et al., 2000; Housner & Griffey, 1985; Swanson et al., 1990). Durch das Studiendesign ist es möglich, dass an drei Messzeitpunkten internes Feedback ausgelöst wird. Die Feedbackgruppe erhält durch den Experten zusätzlich externes Feedback bezogen auf den Ablauf der Stunde. Nach Butler und Winne (1995) ist Feedback besonders effektiv, wenn eine Kombination aus externem und internem Feedback gegeben wird. Evtl. unterschied sich das externe Feedback durch die Expertise des Feedbackgebers von dem internen, eher oberflächlichen Feedback der Probanden. Es wäre denkbar, dass die Probanden der Selbstreflexionsgruppe am dritten Messzeitpunkt bereits Routine im Einsatz der Medien, als Beispiel für oberflächliche Probleme, gefunden hatten und somit den Fokus auf den eigentlichen Unterricht lenken konnten. Bei der Feedbackgruppe könnten unterschiedliche Rückmeldungen des internen und externen Feedbacks dies erschwert haben. Die Hypothese 6.5 kann nicht bestätigt werden.

Reflexion weiterer ausgewählter Ergebnisse:

Die Ergebnisse zur Perspektivenübernahme und zum Literaturbezug lassen erkennen, dass eine signifikante Steigerung erst im dritten Messzeitpunkt möglich ist. Zwischen erstem und zweiten Messzeitpunkt lassen sich keine Unterschiede erkennen. Im Gruppenvergleich der Perspektivenübernahme kann die Kontrollgruppe, im Gegensatz zu den Interventionsgruppen, keine signifikante Steigerung aufweisen. Auch im Literaturbezug schneidet die Kontrollgruppe schlechter ab als die Interventionsgruppen, kann jedoch eine signifikante Steigerung verzeichnen. Allerdings konnten bei beiden Facetten der Reflexionstiefe keine Gruppenunterschiede ermittelt werden.

Die Tendenzen und vor allem das Ergebnis der Unterrichtswirksamkeit lassen eine weitere Auswertung der Daten des Feedbacks und der Methodenwechsel sinnvoll erscheinen. Dieses Design wurde auf Basis der Studie von Anthofer (2017) übernommen, welche unter anderem zum Ziel hatte, eine Entwicklung der Studierenden im Bereich der Schülerversuche vom kochbuchartigen Experimentieren zum offenen Experimentieren zu erreichen. Die Versuche wurden vorgegeben. Auf deren Grundlage wurde auch Feedback erteilt. Somit wäre es interessant zu sehen, inwiefern sich die Methodenwechsel beim Experimentieren, das einen Großteil der Unterrichtsstunde einnahm, unterscheiden. Es ist möglich, dass die Feedbackgruppe mehrmals versucht, die Versuche offener zu gestalten.

In diversen Studien wurde die kognitive Überforderung der Studierenden am Anfang des Berufslebens beschrieben (Barth, 2017; Denner & Gesenhues, 2013; Housner & Griffey, 1985; Swanson et al., 1990). Es wäre evtl. vorstellbar, die Art des Feedbacks zu ändern.

Jahncke et al. (2018) konnten nachweisen, dass Feedback zu schriftlichen Reflexionen über Unterrichtsstunden hohen Einfluss auf die Entwicklung der Reflexionskompetenz hat (Hilzensauer, 2008). Indem die Studierenden aufgefordert werden, über etwaige Punkte der Reflexion erneut nachzudenken und z. B. beschreibende Aspekte bewerten und begründen zu lassen, sowie alternative Handlungsoptionen generieren oder Verbesserungsvorschläge geben zu lassen, würde die individuelle Entwicklung der Studierenden differenziert gefördert werden.

9. Literaturverzeichnis

- Abell, S. K., Appleton, K. & Hanuscin, D. L. (Hrsg.). (2014). *Handbook of research on science education*. New York: Routledge.
- Abell, S. K., Rogers, M. A. P., Hanuscin, D. L., Lee, M. H. & Gagnon, M. J. (2009). Preparing the Next Generation of Science Teacher Educators: A Model for Developing PCK for Teaching Science Teachers. *Journal of Science Teacher Education*, 20(1), 77–93.
- Abels, S. (2011). *LehrerInnen als "Reflective Practitioner"*. Reflexionskompetenz für einen demokratieförderlichen Naturwissenschaftsunterricht. Wiesbaden: Verlag für Sozialwissenschaften.
- Adamson, B. & Walker, E. (2011). Messy collaboration: Learning from a learning study. *Teaching and teacher education*, 27(1), 29–36.
- Al Riyami, T. (2015). Reflection: Is It A Promising Or Spurious Tool For Teachers' Professional Development? *International Journal of Bilingual & Multilingual Teachers of English*, 3(1), 47–58.
- Altrichter, H., Posch, P. & Spann, H. (2018). *Lehrerinnen und Lehrer erforschen ihren Unterricht: Unterrichtsentwicklung und Unterrichtsevaluation durch Aktionsforschung*. Stuttgart: UTB.
- Anderson, J. R. (1988). *Kognitive Psychologie. Eine Einführung*. Heidelberg: Spektrum der Wissenschaft Verlag.
- Anthofer, S. (2017). *Förderung des fachspezifischen Professionswissens von Chemielehramtsstudierenden*. Berlin: Logos Verlag Berlin.
- Argyris, C. & Schön, D. A. (1999). *Die lernende Organisation: Grundlagen, Methode, Praxis*. Stuttgart: Schäffer-Poeschel.
- Artmann, M., Herzmann, P., Hoffmann, M. & Proske, M. (2013). Wissen über Unterricht. Zur Reflexionskompetenz von Studierenden in der ersten Phase der Lehrerbildung. In A. Gehrmann, B. Kranz, S. Pelzmann & A. Reinartz (Hrsg.), *Formation und Transformation der Lehrerbildung. Entwicklungstrends und Forschungsbefunde* (S. 134–150). Bad Heilbrunn: Klinkhardt.
- Aufschnaiter, C. von, Fraij, A. & Kost, D. (2019). Reflexion und Reflexivität in der Lehrerbildung. *Herausforderung Lehrer* innenbildung-Zeitschrift zur Konzeption, Gestaltung und Diskussion*, 2(1), 144–159.
- Backes, A., Sumfleth, E. & Tepner, O. (2012). Test zum experimentell-fachdidaktischen Wissen von Chemielehrkräften. Essen: Unveröffentlichtes Manuskript.
- Bader & Schmidkunz (2002). Das Experiment im Unterricht. In Pfeifer, Lutz & Bader (Hrsg.), *Konkrete Fachdidaktik Chemie* (S. 292–327). München: Oldenburg.
- Bader, R. (2002). Handlungsorientierung in der Berufsbildung. Variantenreiche Ausprägungen. *Die berufsbildende Schule*, 54(3), 71–73.

- Bain, J. D., Ballantyne, R., Packer, J. & Mills, C. (1999). Using journal writing to enhance student teachers' reflectivity during field experience placements. *Teachers and teaching*, 5(1), 51–73.
- Ball, D. L., Hill, H. C. & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, 29(1), 14-17, 20-22, 43-46.
- Ball, D. L., Lubienski, S. T. & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. *Handbook of research on teaching*, 4, 433–456.
- Ballard, K. K. (2006). *Using Van Manen's model to assess levels of reflectivity among preservice physical education teachers*. Texas A&M University, College Station.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
- 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.
- Barnett, R. (1997). *Higher education. A critical business*. Buckingham: Society for research into higher education & Open University Press.
- Barron, B. (2007). Video as a tool to advance understanding of learning and development in peer, family, and other informal learning contexts. In Goldman, Roy, B. Barron & Sharon (Hrsg.), *Video Research in the Learning Sciences* (S. 159–187). London: Routledge.
- Barth, V. L. (2017). *Professionelle Wahrnehmung von Störungen im Unterricht*. Wiesbaden: Springer.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.
- 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. 29–54). Münster: Waxmann.
- Baumert, J., Kunter, M., Brunner, M., Krauss, S., Blum, W. & Neubrand, M. (2004). Mathematikunterricht aus Sicht der PISA-Schülerinnen und -Schüler und ihrer Lehrkräfte. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner & M. Neubrand (Hrsg.), *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (S. 314–354). Münster: Waxmann.
- Baxter, J. A. & Lederman, N. G. (1999). Assessment and measurement of pedagogical content knowledge. In J. Gess-Newsome & N. G. Lederman (Hrsg.), *Examining pedagogical content knowledge* (S. 147–161). Wiesbaden: Springer.
- Beck, E., Baer, M., Guldemann, T., Bischoff, S., Brühwiler, C., Müller, P. et al. (2008). *Adaptive Lehrkompetenz. Analyse und Struktur, Veränderbarkeit und Wirkung handlungssteuernden Lehrerwissens*. Münster: Waxmann.

- Bennewitz, H., Rothland, M. & Terhart, E. (2011). *Handbuch der Forschung zum Lehrerberuf*. Münster: Waxmann.
- Berg, A., Bergendahl, C., Lundberg, B. & Tibell, L. (2003). Benefiting from an open-ended experiment? A comparison of attitudes to, and outcomes of, an expository versus an open-inquiry version of the same experiment. *International Journal of Science Education*, 25(3), 351–372.
- Berliner, D. (1988). *The Development of Expertise in Pedagogy*. Washington, D. C.: American Association of Colleges für Teacher Education.
- Bildungsdirektion Kanton Zürich. (2006). *Klassenführung*. Zugriff am 04.10.2020. Verfügbar unter: www.bildungsdirektion.zh.ch
- Blomberg, G. (2011). *Der Einsatz von Unterrichtsvideos in der universitären Lehrerbildung*. Technische Universität München, München.
- Blömeke, S., Felbrich, A. & Müller, C. (2008). Messung des erziehungswissenschaftlichen Wissens angehender Lehrkräfte. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und-referendare - Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung* (S. 171–195). Münster: Waxmann.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2008). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und-referendare - Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung*. Münster: Waxmann.
- Blömeke, S., Kaiser, G., Lehmann, R., König, J., Döhrmann, M., Buchholtz, C. et al. (2008). *TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich*. Münster: Waxmann.
- Bold, C. (2005). Transforming practice through critical reflection. In C. Bold (Hrsg.), *Supporting Learning and Teaching* (S. 189–202). London: Routledge.
- Bolle, R. (2013). Theoriegeleitete Reflexion in Schulpraxis und Lehrerbildung. In R. Bolle (Hrsg.), *Professionalisierung im Lehramtsstudium. Schulpraktische Kompetenzentwicklung und theoriegeleitete Reflexion* (S. 191–216). Leipzig: Leipziger Universitätsverlag.
- Bönsch, M. (1995). *Differenzierung in Schule und Unterricht: Ansprüche, Formen, Strategien*. München: Ehrenwirth.
- Borko, H., Jacobs, J., Eiteljorg, E. & Pittman, M. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and teacher education*, 24(2), 417–436.
- Borko, H. & Putnam, R. (1996). Learning to Teach. In D. Berliner & R. Calfee (Hrsg.), *Handbook of Educational Psychology* (S. 673–708). New York: MacMillan.

- Borowski, A., Kirschner, S., Liedtke, S. & Fischer, H. E. (2011). Vergleich des Fachwissens von Studierenden, Referendaren und Lehrenden in der Physik. *PhyDid A-Physik und Didaktik in Schule und Hochschule*, 1(10), 1–9.
- Borowski, A., Olszewski, J. & Fischer, H. E. (2010). Fachdidaktisches Wissen von Physikreferendaren. *Der mathematische und naturwissenschaftliche Unterricht*, 63(5), 260–263.
- Bortz, J. & Döring, N. (2007). *Forschungsmethoden und Evaluation für Human-und Sozialwissenschaftler: Limitierte Sonderausgabe*. Wiesbaden: Springer.
- Bortz, J. & Schuster, C. (2011). *Statistik für Human-und Sozialwissenschaftler: Limitierte Sonderausgabe*. Wiesbaden: Springer.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (2012). *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann Verlag.
- Bosse, D. (2012). Zur Situation der Lehrerbildung in Deutschland. In D. Bosse, L. Criblez & T. Hascher (Hrsg.), *Reform der Lehrerbildung in Deutschland, Österreich und der Schweiz* (Theorie und Praxis der Schulpädagogik, Bd. 4, S. 11–28). Immenhausen bei Kassel: Prolog Verlag.
- Boud, D., Keogh, R. & Walker, D. (1985). Promoting reflection in learning: A model. In D. Boud, R. Keogh & D. Walker (Hrsg.), *Reflection: turning experience into learning* (S. 18–40). London: Kogan Page.
- Boyd, W. L. & Reese, J. P. (2006). Great expectations: The impact of the national board for professional teaching standards. *Education Next*, 6(2), 50–58.
- Bradburn, N. M., Sudman, S. & Wansink, B. (2004). *Asking questions: the definitive guide to questionnaire design--for market research, political polls, and social and health questionnaires*. San Francisco: John Wiley & Sons.
- Bräuer, G. (2009). Reflecting the practice of foreign language learning in portfolios. *German as a Foreign Language-Journal*, (2-3), 148–166.
- Bräuer, G. (2016). *Das Portfolio als Reflexionsmedium für Lehrende und Studierende*. Stuttgart: UTB.
- Bräuer, G. & Schindler, K. (2011). *Schreibarrangements für Schule, Hochschule, Beruf*. Stuttgart: Ernst Klett Sprachen GmbH.
- Brendel, N. (2017). *Reflexives Denken im Geographieunterricht: eine empirische Studie zur Bestimmung von Schülerreflexion mithilfe von Weblogs im Kontext Globalen Lernens*. Münster: Waxmann Verlag.
- Bromme, R. (2014). *Der Lehrer als Experte: Zur Psychologie des professionellen Wissens*. Münster: Waxmann Verlag.
- Brophy (Hrsg.). (2003). *Using video in teacher education. (Advances in Research on Teaching)*. Bingley: Emerald Group Publishing Limited.

- Brophy, J., Good, T. & Wittrock, M. C. (1986). *Handbook of research on teaching*. New York: MacMillan.
- Brouër, B. & Gläser-Zikuda, M. (2010). Förderung selbstregulativer Fähigkeiten im Kontext selbstorganisierten Lernens. *Zeitschrift für Berufs- und Wirtschaftspädagogik, Beiheft Bd. 23*, 123–136.
- Brouwer, N. & Korthagen, F. (2005). Can teacher education make a difference? *American Educational Research Journal*, 42(1), 153–224.
- Brunner, M., Kunter, M., Krauss, S., Klusmann, U., Baumert, J., Blum, W. et al. (2006). Die professionelle Kompetenz von Mathematiklehrkräften. Konzeptualisierung, Erfassung und Bedeutung für den Unterricht. Eine Zwischenbilanz des COACTIV-Projekts. *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms*, 54–82.
- Bühl, A. (2014). *SPSS 22. Einführung in die moderne Datenanalyse*. Hallbergmoos: Pearson Deutschland GmbH.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. Hallbergmoos: Pearson Deutschland GmbH.
- Bühner, M. & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. Hallbergmoos: Pearson Deutschland GmbH.
- Buhren, C. G. (2015). Feedback - Definitionen und Differenzierungen. In C. G. Buhren (Hrsg.), *Handbuch Feedback in der Schule* (S. 11–30). Weinheim und Basel: Beltz.
- Butler, D. L. & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of educational research*, 65(3), 245–281.
- Byers, A., Koba, S., Sherman, G., Scheppke, J. & Bolus, R. (2011). Developing a web-based mechanism for assessing teacher science content knowledge. *Journal of Science Teacher Education*, 22(3), 273–289.
- Cajkler, W., Wood, P., Norton, J., Pedder, D. & Xu, H. (2015). Teacher perspectives about lesson study in secondary school departments: a collaborative vehicle for professional learning and practice development. *Research Papers in Education*, 30(2), 192–213.
- Calderhead, J. (1989). Reflective teaching and teacher education. *Teaching and teacher education*, 5(1), 43–51.
- Calderhead, J. (1996). Teachers: Beliefs and knowledge. In D. Berliner & R. Calfee (Hrsg.), *Handbook of Educational Psychology* (S. 709–725). New York: MacMillan.
- Carlson, J. & Daehler, K. R. (2019). The Refined Consensus Model of Pedagogical Content Knowledge in Science Education. In A. Hume, R. Cooper & A. Borowski (Hrsg.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (S. 77–92). Wiesbaden: Springer.
- Carrington, S. & Selva, G. (2010). Critical social theory and transformative learning: Evidence in pre-service teachers' service-learning reflection logs. *Higher Education Research & Development*, 29(1), 45–57.

- Carter, K., Cushing, K., Sabers, D., Stein, P. & Berliner, D. (1988). Expert-novice differences in perceiving and processing visual classroom information. *Journal of teacher education*, 39(3), 25–31.
- Champagne, A. B. & Klopfer, L. E. (1977). A sixty-year perspective on three issues in science education: I. Whose ideas. Are Dominant? II. Representation of women. III. Reflective thinking and problem solving. *Science Education*, 61(4), 431–452.
- Chandler, P., Robinson, W. P. & Noyes, P. (1990). Is initial training nurturing proactive teachers? *Educational Research*, 32(2), 130–139.
- Cheng, E. C. K. & Lo. (2013). The approach of learning study: Its origin and implications, 1–28. Zugriff am 04.10.2020. Verfügbar unter: <https://www.oecd-ilibrary.org/content/paper/5k3wjp0s959p-en>
- Çimer, A., Çimer, S. O. & Vekli, G. S. (2013). How does reflection help teachers to become effective teachers. *International Journal of Educational Research*, 1(4), 133–149.
- Clausen, M., Reusser, K. & Klieme, E. (2003). Unterrichtsqualität auf der Basis hochinferenter Unterrichtsbeurteilungen: Ein instruktionspsychologischer Vergleich zwischen Deutschland und der deutschsprachigen Schweiz. *Unterrichtswissenschaft*, 31(2), 122–141.
- Clauss, Finze, F. & Partzsch, L. (2004). Statistik-Grundlagen für Soziologen, Pädagogen, Psychologen und Mediziner. *Frankfurt a. M.: Harri Deutsch*.
- Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational researcher*, 23(7), 13–20.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155–159.
- Collin, S., Karsenti, T. & Komis, V. (2013). Reflective practice in initial teacher training: Critiques and perspectives. *Reflective practice*, 14(1), 104–117.
- Combe & Kolb, D. (2004). Lehrerprofessionalität: Wissen, Können, Handeln. In W. Helsper & J. Böhme (Hrsg.), *Handbuch der Schulforschung* (S. 833–851). Wiesbaden: Verlag für Sozialwissenschaften.
- Combe, A. (1997). Pädagogische Professionalität, Hermeneutik und Lehrerbildung. Am Beispiel der Berufsbelastung von Grundschullehrern. In A. Combe & W. Helsper (Hrsg.), *Pädagogische Professionalität. Untersuchungen zum Typus pädagogischen Handelns* (S. 501–520). Frankfurt am Main: Suhrkamp.
- Converse, J. M. & Presser, S. (1986). *Survey questions: Handcrafting the standardized questionnaire* (Bd. 63). Thousand Oaks, California: Sage.
- Copeland, W. D., Birmingham, C., La Cruz, E. de & Lewin, B. (1993). The reflective practitioner in teaching: Toward a research agenda. *Teaching and teacher education*, 9(4), 347–359.

- Cornford, I. R. (2002). Reflective teaching: Empirical research findings and some implications for teacher education. *Journal of Vocational education and Training*, 54(2), 219–236.
- Corno, L. & Snow, R. E. (1986). Adapting teaching to individual differences among learners. *Handbook of research on teaching*, 3, 605–629.
- Daheim, H. (1992). Zum Stand der Professionssoziologie. Rekonstruktion machttheoretischer Modelle der Profession. In Dewe, Ferchhoff & Radtke (Hrsg.), *Erziehen als Profession. Zur Logik professionellen Handelns in pädagogischen Feldern* (S. 21–35). Wiesbaden: Springer.
- Danielson, L. (2008). Making reflective practice more concrete through reflective decision making. *The Educational Forum*, 72(2), 129–137.
- Darling-Hammond, L. (2012). *Powerful teacher education. Lessons from exemplary programs*. Hoboken, New Jersey: John Wiley & Sons.
- Darling-Hammond, L., Berry, B. & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational evaluation and policy analysis*, 23(1), 57–77.
- Den Brok, P., Brekelmans, M. & Wubbels, T. (2004). Interpersonal teacher behaviour and student outcomes. *School effectiveness and school improvement*, 15(3-4), 407–442.
- Denner, L. & Gesenhues, D. (2013). Professionalisierungsprozesse im Lehramtsstudium. eine explorative Studie zu Analyse, Interpretation und Handlungsoption. In R. Bolle (Hrsg.), *Professionalisierung im Lehramtsstudium. Schulpraktische Kompetenzentwicklung und theoriegeleitete Reflexion* (S. 59–120). Leipzig: Leipziger Universitätsverlag.
- Deutscher Bildungsrat. (1970). *Empfehlungen der Bildungskommission. Strukturplan für das Bildungswesen: Verabschiedet auf der 27. Sitzung der Bildungskommission am 13. Februar 1970*. Stuttgart: Klett.
- Dewey, J. (1910). *How we think*. Boston: D.C. Heath & Co.
- Dewey, J. (1933). *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. Lexington, Massachusetts: Heath.
- Dohrmann, R. & Nordmeier, V. (2018). Professionalität im Lehr-Lern-Labor anbahnen-Ergebnisse zu verschiedenen Facetten von Reflexion und Selbstwirksamkeitserwartungen. *PhyDid B-Didaktik der Physik-Beiträge zur DPG-Frühjahrstagung*, 1, 73–80.
- Dreyfus, H. L., Dreyfus, S. E. & Athanasiou, T. (2000). *Mind over machine*. New York: Simon and Schuster.
- Dymoke, S. & Harrison, J. (2008). *Reflective teaching and learning*. Thousand Oaks, California: SAGE Publications Ltd.

- Eickhorst, A. (2011). Das Unterrichtsverständnis der empirischen Lehr-Lern-Forschung. In Meseth, Proske & Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 50–66). München: Klinkhardt.
- Eid, M. & Schmidt, K. (2014). *Testtheorie und Testkonstruktion*. Göttingen: Hogrefe Verlag.
- Einsiedler, W. (2002). Das Konzept Unterrichtsqualität. *Unterrichtswissenschaft*, 30(3), 194–196.
- El-Dib, M. A. B. (2007). Levels of reflection in action research. An overview and an assessment tool. *Teaching and teacher education*, 23(1), 24–35.
- Emmer & Gerwels (2006). Classroom management in middle and high school classrooms. In Evertson & Weinstein (Hrsg.), *Handbook of Classroom Management. Research, practice and contemporary issues* (S. 407–437). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Estrada, F. F. & Rahman, H. M.A. (2014). Reflective journal writing as an approach to enhancing students' learning experience. *Brunei Darussalam Journal of Technology and Commerce*, 8(1), 22–35.
- Evertson, C. M. & Weinstein, C. S. (2013). *Handbook of classroom management: Research, practice, and contemporary issues*. London: Routledge.
- Farrell, T. S. C. (2014). "Teacher You Are Stupid!"-Cultivating a Reflective Disposition. *The Electronic Journal for English as a Second Language*, 18(3), 1-10.
- Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160.
- Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, (39), 175–191.
- Feiks, D. & Krauß, E. (2001). *Professionell handeln, erziehen, unterrichten: Konzepte und Beispiele*. Stuttgart: E. Klett.
- Fend, H. (1981). *Theorie der Schule*. München: Urban & Schwarzenbeck.
- Fengler, J. (2009). *Feedback geben. Strategien und Übungen*. Weinheim: Beltz.
- Fernandez, Cannon & Chokshi. (2003). A US–Japan lesson study collaboration reveals critical lenses for examining practice. *Teaching and teacher education*, 19(2), 171–185.
- Ferri, R. B. & Blum, W. (2018). *Lehrerkompetenzen zum Unterrichten mathematischer Modellierung: Konzepte und Transfer*. Wiesbaden: Springer.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. New York: Sage.

- Fischer, C., Rieck, K. & Lobemeier, K. R. (2008). Mit Logbüchern dokumentieren und reflektieren. Das Beispiel SINUS-Transfer Grundschule. In Lankes (Hrsg.), *Pädagogische Professionalität als Gegenstand empirischer Forschung* (S. 73–85). Münster: Waxmann.
- Fischer, H. E. & Sumfleth, E. (2013). *nwu-essen: 10 Jahre Essener Forschung zum naturwissenschaftlichen Unterricht*. Berlin: Logos Verlag Berlin.
- Fischler, H. (2008). Physikdidaktisches Wissen und Handlungskompetenz. *ZfDN*, (14), 27–42.
- Fisseni, H. J. (2004). *Lehrbuch der psychologischen Diagnostik: mit Hinweisen zur Intervention*. Göttingen: Hogrefe Verlag.
- Fraser (1998). Science Learning Environments: Assessment, Effects and Determinants. In Fraser & Tobin (Hrsg.), *International handbook of science education - part one* (S. 527–564). London: Kluwer Academic Publishers.
- Fraser, B. J. (1991). Two decades of classroom environment research. In Fraser & Walberg (Hrsg.), *Educational environments: Evaluation, antecedents and consequences*. Headington: Pergamon Press.
- Frey, A. (2004). Die Kompetenzstruktur von Studierenden des Lehrerberufs. *Zeitschrift für Pädagogik*, 50(6), 903–925.
- Frey, A. (2006). Methoden und Instrumente zur Diagnose beruflicher Kompetenzen von Lehrkräften—eine erste Standortbestimmung zu bereits publizierten Instrumenten. *Zeitschrift für Pädagogik*, 51, 30–46.
- Frey, A., Jäger, R. S. & Renold, U. (2003). *Kompetenzmessung: Sichtweisen und Methoden zur Erfassung und Bewertung von beruflichen Kompetenzen*. Landau: Verein Empirische Pädagogik.
- Frey, A. & Jung, C. (2011). Kompetenzmodelle und Standards in Lehrerbildung und Lehrerberuf. In Terhart, Bennewitz & Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 540–572). Münster: Waxmann.
- Frey, A., Taskinen, P., Schütte, K., Prenzel, M., Artelt, C., Baumert, J. et al. (2009). *PISA 2006. Skalenhandbuch. Die Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Früh, W. (2015). *Inhaltsanalyse: Theorie und Praxis*. Stuttgart: UTB.
- Gärtner, H. (2007). *Unterrichtsmonitoring*. Münster: Waxmann.
- Gehrmann, A., Hericks, U. & Lüders, M. (Hrsg.). (2010). *Bildungsstandards und Kompetenzmodelle. Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht*. Bad Heilbrunn: Klinkhardt.
- George, D. & Mallery, M. (2002). *SPSS for Windows step by step. A simple guide and reference*. London: Routledge.

- Geppert, K. & Preuss, E. (1978). *Differenzierender Unterricht konkret: Analyse, Planung und Gestaltung; ein Modell zur Reform des Primarbereichs*. München: Klinkhardt.
- Gess-Newsome, J. (2015). A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK Summit. In A. Berry, P. Friedrichsen & J. Loughran (Hrsg.), *Re-examining Pedagogical Content Knowledge in Science Education* (S. 38–52). New York: Routledge.
- Gess-Newsome, J., Taylor, J. A., Carlson, J., Gardner, A. L., Wilson, C. D. & Am Stuhlsatz, M. (2017). Teacher pedagogical content knowledge, practice, and student achievement. *International Journal of Science Education*, 41(7), 944–963.
- Girmes, R. (2006). Lehrprofessionalität in einer demokratischen Gesellschaft. *Zeitschrift für Pädagogik*, 52(51), 14–29.
- Gold, B., Hellermann, C. & Holodynski, M. (2017). Effekte videobasierter Trainings zur Förderung der Selbstwirksamkeitsüberzeugungen über Klassenführung im Grundschulunterricht. *Zeitschrift für Erziehungswissenschaft*, 20(1), 115–136.
- Gore, J. M. (1987). Reflecting on reflective teaching. *Journal of teacher education*, 38(2), 33–39.
- Gore, J. M. & Zeichner, K. (1991). Action research and reflective teaching in preservice teacher education: A case study from the United States. *Teaching and teacher education*, 7(2), 119–136.
- Greenaway, R. (2008). A view into the future: the value of other ways of learning and development. In P. Becker & J. Schirp (Hrsg.), *Other Ways of Learning. The European Institute for Outdoor Adventure Education and Experimental Learning 1996–2006* (S. 347–367). Marburg: bsj.
- Großebrahm, N. (2013). *Elemente fachdidaktischen Wissens in der universitären Ausbildung angehender Chemielehrkräfte. Ein Beitrag zur Standardentwicklung*. Dissertation. Duisburg, Essen, Universität, Duisburg, Essen.
- Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education. (Professional Development and Practice Series)*. Columbia: Teachers College Press.
- Gruber, H. (2008). Lernen und Wissenserwerb. In W. Schneider, M. Hasselhorn & J. Bengel (Hrsg.), *Handbuch der pädagogischen Psychologie* (S. 95–104). Göttingen: Hogrefe.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung*. Münster: Waxmann.
- Gruschka, A. (2011). Der empirische Blick auf das Unterrichten als pädagogischer Prozess. In Meseth, Proske & Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 130–144). München: Klinkhardt.

- Guardiera, P., Podlich, C. & Reimer, A. (2018). Zur Förderung von Reflexionskompetenz in der LehrerInnenbildung. In N. Ukley & B. Gröben (Hrsg.), *Forschendes Lernen im Praxissemester. Begründungen, Befunde und Beispiele aus dem Fach Sport* (S. 231–248). Wiesbaden: Springer.
- Guldimann, T., Bischoff, S. & Brühwiler, C. (2005). Adaptive Lehrkompetenz. Analyse von Struktur, Veränderbarkeit und Wirkung von handlungssteuerndem Lehrwissen. *Beiträge zur Lehrerbildung*, 25(2), 262–269.
- Gurney, P. (2007). Five factors for effective teaching. *New Zealand Journal of Teachers' Work*, 4(2), 89–98.
- Haag, L. & Götz, T. (2019). Was wissen wir über Feedback? Synopse des aktuellen Forschungsstandes. *Friedrich Jahresheft*, XXXVII, 14–19.
- Haas, A. (1998). *Unterrichtsplanung im Alltag: Eine empirische Untersuchung zum Planungs Handeln von Hauptschul-, Realschul- und Gymnasiallehrern*. Berlin: Roderer.
- Haase, J. (2018). *Selbstbestimmtes Lernen im naturwissenschaftlichen Sachunterricht: eine empirische Interventionsstudie mit Fokus auf Feedback und Kompetenzerleben*. Berlin: Logos Verlag Berlin.
- Hänssig, A. (2010). Portfolio-Arbeit in den Schulpraktischen Studien. Ein Praxisbeispiel. In K. Liebsch (Hrsg.), *Reflexion und Intervention. Zur Theorie und Praxis Schulpraktischer Studien* (S. 141–180). Baltmannsweiler: Schneider Verlag Hohengehren.
- Harlen, W. (1999). *Effective Teaching of Science. A Review of Research*. Edinburgh: Scottish Council for Research in Education.
- Harrington, H. L., Quinn-Leering, K. & Hodson, L. (1996). Written case analyses and critical reflection. *Teaching and teacher education*, 12(1), 25–37.
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York: Routledge.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81–112.
- Hattie, J. & Zierer, K. (2016). *"Kenne deinen Einfluss!". "Visible Learning" für die Unterrichtspraxis*. Hohengehren: Schneider Verlag.
- Hatton, N. & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and teacher education*, 11(1), 33–49.
- Haußer, K., Mayring, P. & Strehmel, P. (1982). Praktische Probleme bei der Inhaltsanalyse offen erhobener Kognitionen, diskutiert am Beispiel der Variablen Berufsinteresse arbeitsloser Lehrer. In H.-D. Dann, W. Humpert & Tennstädt. K.-C. (Hrsg.),

- Analyse und Modifikation subjektiver Theorien von Lehrern. Analyse und Modifikation subjektiver Theorien von Lehrern, Ergebnissen und Perspektiven eines Kolloquiums.* Konstanz: Universität Konstanz.
- Hayes, A. F. & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, 1(1), 77–89.
- Heber, R. (2018). *Erfassung und Entwicklung von Feedback-Metawissen bei Chemie-lehramtsstudierenden*. Zulassungsarbeit. Universität Regensburg, Regensburg.
- Heller, K. & Perleth, C. (2000). *KFT 4-12+ R-Kognitiver Fähigkeiten Test für 4. bis 12 [The Cognitive Abilities Test for grades 4 to 12]*. Klassen. Göttingen: Beltz.
- Helmke, A. (2007a). Lernprozesse anregen und steuern: was wissen wir über Klarheit und Strukturiertheit? *Pädagogik*, 59(6), 44–47.
- Helmke, A. (2007b). *Unterrichtsqualität: Erfassen, bewerten, verbessern*. Stuttgart: Klett.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett/Kallmeyer.
- Helmke, A., Goellner, R., Kleinbub, I., Schrader, F. W. & Wagner, W. (2008). Standards–Motor oder Bremse der Unterrichtsentwicklung. In Bundesarbeitskreis der Seminar- und Fachleiter/innen e. V. (Hrsg.), *Seminar–Lehrerbildung und Schule* (Bd. 1, S. 62–75).
- Helmke, A. & Klieme, E. (2008). *Unterricht und Entwicklung sprachlicher Kompetenzen*.
- Helmke, A. & Schrader, F.-W. (2010). Merkmale der Unterrichtsqualität: Potenzial, Reichweite und Grenzen. *Qualitätssicherung im Bildungswesen. Auftrag und Anspruch der bayerischen Qualitätsagentur*, 1, 17–47.
- Helsper, W. & Tippelt, R. (2011). *Pädagogische Professionalität*. Weinheim: Beltz.
- Henderson, J. G. (1992). *Reflective teaching: Becoming an inquiring educator*. London: MacMillan.
- Hericks, U. (2006). *Professionalisierung als Entwicklungsaufgabe*. Wiesbaden: Springer.
- Herzig, B., Grafe, S. & Reinhold, P. Reflexives Lernen mit digitalen Videos. Zur Auseinandersetzung mit dem Theorie-Praxisverhältnis in der Lehrerbildung. In M. Welzel & H. Stadler (Hrsg.), *Nimm doch mal die Kamera! Zur Nutzung von Videos in der Lehrerbildung* (S. 45–64). Münster: Waxmann.
- Herzog, W. (2003). Zeitgemäße Erziehung: die Konstruktion pädagogischer Wirklichkeit. *Zeitschrift für Pädagogik*, 49(3), 619–622.
- Hill, H. C., Ball, D. L. & Schilling, S. G. (2008). Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students. *Journal for Research in Mathematics Education*, 39(4), 372–400.

- Hilzensauer, W. (2008). Theoretische Zugänge und Methoden zur Reflexion des Lernens. Ein Diskussionsbeitrag. *bildungsforschung*, 5(2), 1–18.
- Hofstein, A. & Lunetta, V. N. (1982). The role of the laboratory in science teaching: Neglected aspects of research. *Review of educational research*, 52(2), 201–217.
- Holmqvist, M. (2011). Teachers' learning in a learning study. *Instructional Science*, 39(4), 497–511.
- Holsti, O. R. (1969). *Content analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley.
- Hoppe-Graff, S., Schroeter, R. & Flagmeyer, D. (2008). Universitäre Lehrerbildung auf dem Prüfstand: Wie beurteilen Referendare das Theorie-Praxis-Problem. *Empirische Pädagogik*, 22(3), 353–381.
- Hosenfeld, A. (2010). *Führt Unterrichtsrückmeldung zu Unterrichtsentwicklung?* Münster: Waxmann.
- Housner, L. D. & Griffey, D. C. (1985). Teacher cognition: Differences in planning and interactive decision making between experienced and inexperienced teachers. *Research Quarterly for Exercise and Sport*, 56(1), 45–53.
- Hsieh, S.-W., Jang, Y.-R., Hwang, G.-J. & Chen, N.-S. (2011). Effects of teaching and learning styles on students' reflection levels for ubiquitous learning. *Computers & Education*, 57(1), 1194–1201.
- Hugener, I., Krammer, K. & Pauli, C. (2008). Kompetenzen der Lehrpersonen im Umgang mit Heterogenität: Differenzierungsmaßnahmen im Mathematikunterricht. In M. Gläser-Zikuda & J. Seifried (Hrsg.), *Lehrerexpertise. Analyse und Bedeutung unterrichtlichen Handelns* (S. 47–66). Münster: Waxmann.
- Hullfish, H. G. & Smith, P. G. (1961). *Reflective thinking: the method of education*. New York: Dodd, Mead.
- Hume, A. (2009). Promoting higher levels of reflective writing in student journals. *Higher Education Research & Development*, 28(3), 247–260.
- Hung, I.-C., Yang, X.-J., Fang, W.-C., Hwang, G.-J. & Chen, N.-S. (2014). A context-aware video prompt approach to improving students' in-field reflection levels. *Computers & Education*, 70, 80–91.
- Jaehnig, W. & Miller, M. L. (2007). Feedback types in programmed instruction: A systematic review. *The psychological record*, 57(2), 219–232.
- Jahncke, H., Berding, F., Porath, J. & Magh, K. (2018). Einfluss von Feedback auf die (Selbst-) Reflexion von Lehramtsstudierenden. *Beiträge zu Praxis, Praxisforschung und Forschung Jahrgang 2018*, 505–530.
- Janich, P. (1995). Experiment. *Enzyklopädie Philosophie und Wissenschaftstheorie*, (1), 621–622.

- Janssen, J. & Laatz, W. (2013). *Statistische Datenanalyse mit SPSS für Windows: eine anwendungsorientierte Einführung in das Basissystem Version 8 und das Modul Exakte Tests*. Wiesbaden: Springer.
- Jarvis, P. (1995). *Adult and continuing education: Theory and practice*. London: Psychology Press.
- Jay, J. K. & Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and teacher education*, 18(1), 73–85.
- Jungmeister, A. (2016). *Reflexion und Innovation im Forschungsprozess*. Münster: LIT Verlag.
- Käpylä, M., Heikkinen, J.-P. & Asunta, T. (2009). Influence of content knowledge on pedagogical content knowledge: The case of teaching photosynthesis and plant growth. *International Journal of Science Education*, 31(10), 1395–1415.
- Kelava, A. & Moosbrugger, H. (2012). Deskriptivstatistische Evaluation von Items (Itemanalyse) und Testwertverteilungen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 75–102). Wiesbaden: Springer.
- Keller, S. & Bräuer, G. (2013). Elektronische Portfolios als Katalysatoren für Studium und Lehre. In B. Koch-Priewe, T. Leonhard, A. Pineker & J. Störtländer (Hrsg.), *Portfolio in der LehrerInnenbildung. Konzepte und empirische Befunde*. (S. 265–275). Bad Heilbrunn: Klinkhardt.
- Kember, D., Jones, A., Loke, A., McKay, J., Sinclair, K., Tse, H. et al. (1999). Determining the level of reflective thinking from students' written journals using a coding scheme based on the work of Mezirow. *International Journal of Lifelong Education*, 18(1), 18–30.
- Kember, D., Leung, D. Y. P., Jones, A., Loke, A. Y., McKay, J., Sinclair, K. et al. (2000). Development of a questionnaire to measure the level of reflective thinking. *Assessment & evaluation in Higher education*, 25(4), 381–395.
- Kempin, M., Kulgemeyer, C. & Schecker, H. (2019). Erste Einblicke in die Entwicklung der Reflexionsfähigkeit von Physiklehramtsstudierenden im Praxissemester. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe* (S. 357–360). Regensburg: Universität Regensburg.
- King, P. M. & Kitchener, K. S. (1994). *Developing Reflective Judgment: Understanding and Promoting Intellectual Growth and Critical Thinking in Adolescents and Adults*. Jossey-Bass Higher and Adult Education Series and Jossey-Bass Social and Behavioral Science Series. San Francisco: Jossey-Bass.
- Kipnis, M. & Hofstein, A. (2008). The inquiry laboratory as a source for development of metacognitive skills. *International Journal of Science and Mathematics Education*, 6(3), 601–627.
- Kirschner, S. (2013). *Modellierung und Analyse des Professionswissens von Physiklehrkräften*. Berlin: Logos Verlag Berlin.

- Kitchenham, A. (2008). The evolution of John Mezirow's transformative learning theory. *Journal of transformative education*, 6(2), 104–123.
- Kleinknecht, M. & Poschinski, N. (2014). Eigene und fremde Videos in der Lehrerfortbildung. Eine Fallanalyse zu kognitiven und emotionalen Prozessen beim Beobachten zweier unterschiedlicher Videotypen. *Zeitschrift für Pädagogik*, 60(3), 471–490.
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. In M. Prenzel, L. Allolio-Näcke & Prenzel-Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 127–146). Münster: Waxmann.
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222–237.
- Klusmann, U., Kunter, M., Trautwein, U. & Baumert, J. (2006). Lehrerbelastung und Unterrichtsqualität aus der Perspektive von Lehrenden und Lernenden. *Zeitschrift für pädagogische Psychologie*, 20(3), 161–173.
- KMK (Hrsg.). (2004). *Standards für die Lehrerbildung: Bildungswissenschaften. (Beschluss der Kultusministerkonferenz vom 16.12.2004)*. Zugriff am 04.10.2020. Verfügbar unter: http://www.kmk.org/fileadmin/veroeffentlichungen_beschlusse/2004/2004_12_16-Standards-Lehrerbildung.pdf
- KMK. (2008). Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung. *Beschluss der KMK vom, 16, 1–135*. Verfügbar unter: https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschlusse/2008/2008_10_16-Fachprofile-Lehrerbildung.pdf
- Koch-Priewe, B. (2002). Der routinierte Umgang mit Neuem. Wie die Professionalisierung von JunglehrerInnen gelingen kann. *Jahrbuch für Lehrerforschung und Bildungsarbeit*, 3, 311–324.
- Kohler, B. & Wacker, A. (2013). Das Angebots-Nutzungs-Modell: Überlegungen zu Chancen und Grenzen des derzeit prominentesten Wirkmodells der Schul- und Unterrichtsforschung. *Die Deutsche Schule*, 105(3), 241–257.
- Kolb, D. (1984). *Experiential education: Experience as the source of learning and development*. Upper Saddle River, New Jersey: Pearson Education.
- Kolb, D. (2014). *Experiential learning: Experience as the source of learning and development*. Upper Saddle River, New Jersey: Pearson Education.
- König, J. (2014). Designing an international instrument to assess teachers' general pedagogical knowledge (GPK): Review of studies, considerations, and recommendations. In OECD (Hrsg.), *Technical paper prepared for the OECD innovative Teaching for Effective Learning (ITEL). Phase II, Project: A Survey to Profile the Pedagogical Knowledge in the Teaching Profession*. Paris: OECD.
- König, J. & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 12(3), 499–527.

- König, J. & Seifert, A. (2012a). Der Erwerb von pädagogischem Professionswissen: Ziele, Design und zentrale Ergebnisse der LEK-Studie. In J. König & A. Seifert (Hrsg.), *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerbildung* (S. 7–31). Münster: Waxmann Verlag.
- König, J. & Seifert, A. (Hrsg.). (2012b). *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerbildung*. Münster: Waxmann Verlag.
- Korthagen, F. & Kessels, J. (1999). Linking theory and practice: Changing the pedagogy of teacher education. *Educational researcher*, 28(4), 4–17.
- Korthagen, F., Kessels, J., Koster, B., Lagerwerf, B. & Wubbels, T. (2002a). Eine Reflexion über Reflexion. In F. Korthagen, J. Kessels, B. Koster, B. Lagerwerf & T. Wubbels (Hrsg.), *Schulwirklichkeit und Lehrerbildung: Reflexion der Lehrertätigkeit* (S. 55–73). Hamburg: EB-Verlag.
- Korthagen, F., Kessels, J., Koster, B., Lagerwerf, B. & Wubbels, T. (Hrsg.). (2002b). *Schulwirklichkeit und Lehrerbildung: Reflexion der Lehrertätigkeit*. Hamburg: EB-Verlag.
- Korthagen, F. & Vasalos, A. (2005). Levels in reflection: Core reflection as a means to enhance professional growth. *Teachers and teaching*, 11(1), 47–71.
- Korthagen, F. & Vasalos, A. (Hrsg.). (2009). *From reflection to presence and mindfulness: 30 years of developments concerning the concept of reflection in teacher education*. New York: Citeseer.
- Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York: Holt, Rinehart and Winston.
- Krammer, K. & Reusser, K. (2005). Unterrichtsvideos als Medium der Aus- und Weiterbildung von Lehrpersonen. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 23(1), 35–50.
- Krammer, K., Schnetzler, C. L., Pauli, C., Reusser, K., Ratzka, N., Lipowsky, F. et al. (2010). Unterrichtsvideos in der Lehrerfortbildung: Überblick über Konzeption und Ergebnisse einer einjährigen netzgestützten Fortbildungsveranstaltung. In F. H. Müller, A. Eichenberger, M. Lüders & J. Mayr (Hrsg.), *Lehrerinnen und Lehrer lernen: Konzepte und Befunde zur Lehrerfortbildung* (S. 227–243). Münster: Waxmann.
- Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M. et al. (2008). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und -Lehrer im Rahmen der COACTIV-Studie. *Journal für Mathematik-Didaktik*, 29(3-4), 233–258.
- Krippendorff, K. (1980). *Content Analysis. An introduction to its methodology*. London: Sage.
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. London: Sage publications.

- Kulik, J. A. & Kulik, C.-L. C. (1988). Timing of feedback and verbal learning. *Review of educational research*, 58(1), 79–97.
- Kunina-Habenicht, O., Lohse-Bossenz, H., Kunter, M., Dicke, T., Förster, D., Gößling, J. et al. (2012). Welche bildungswissenschaftlichen Inhalte sind wichtig in der Lehrerbildung? *Zeitschrift für Erziehungswissenschaft*, 15(4), 649–682.
- Kuno, H. (2011). Conceptualizing lesson study as change management recipe. In AEO Nazarbayev Intellectual School, Center of Excellence, the University of Cambridge (Hrsg.), *International Scientifically-Practical Conference: Teacher Professional Development: Traditions and Changes* (S. 4–12).
- Kunter, M. & Baumert, J. Linking TIMSS to research on learning and instruction: A re-analysis of the German TIMSS and TIMSS video data. In S. Howie & T. Plomp (Hrsg.), *Learning mathematics and science. Lesson learned from TIMSS* (S. 335–351). London: Routledge.
- Kunter, M., Baumert, J. & Blum, W. (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Waxmann Verlag.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Kunter, M., Klusmann, U. & Baumert, J. (2009). Professionelle Kompetenz von Mathematiklehrkräften: Das COACTIV-Modell. In O. Zlatkin-Troitschanskaia, Beck, K., Sembill, D., R. Nickolaus & R. Mulder (Hrsg.), *Lehrprofessionalität - Bedingungen, Genese, Wirkungen und ihre Messung* (S. 153–165). Weinheim: Beltz.
- Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 85–113). Münster: Waxmann.
- Labudde, P. & Adamina, M. (2008). HarmoS Naturwissenschaften: Impulse für den naturwissenschaftlichen Unterricht von morgen. *Beiträge zur Lehrerinnen-und Lehrerbildung*, 26(3), 351–360.
- Lamprecht, J. (2011). *Ausbildungswege und Komponenten professioneller Handlungskompetenz: Vergleich von Quereinsteigern mit Lehramtsabsolventen für Gymnasien im Fach Physik*. Berlin: Logos Verlag Berlin.
- Le Fevre, D. M. (2003). Designing for teacher learning: Video-based curriculum design. In Brophy (Hrsg.), *Using video in teacher education. (Advances in Research on Teaching)* (Bd. 10, S. 235–258). Bingley: Emerald Group Publishing Limited.
- Lee, H.-J. (2005). Understanding and assessing preservice teachers' reflective thinking. *Teaching and teacher education*, 21(6), 699–715.
- Lee, I. (2008). Fostering preservice reflection through response journals. *Teacher Education Quarterly*, 35(1), 117–139.

- Leijen, Ä., Allas, R., Toom, A., Husu, J., Marcos, J.-J. M., Meijer, P. et al. (2014). Guided reflection for supporting the development of student teachers' practical knowledge. *Procedia-Social and Behavioral Sciences*, 112, 314–322.
- Lenzner, T., Kaczmirek, L. & Lenzner, A. (2010). Cognitive burden of survey questions and response times: A psycholinguistic experiment. *Applied cognitive psychology*, 24(7), 1003–1020.
- Leonhard, T. (2008). *Professionalisierung in der Lehrerbildung: eine explorative Studie zur Entwicklung professioneller Kompetenzen in der Lehrererstausbildung*. Berlin: Logos-Verlag.
- Leonhard, T., Nagel, N., Rihm, T., Strittmatter-Haubold, V. & Wengert-Richter, P. (2010). Zur Entwicklung von Reflexionskompetenz bei Lehramtsstudierenden. In A. Gehrman, U. Hericks & M. Lüders (Hrsg.), *Bildungsstandards und Kompetenzmodelle. Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht* (S. 111–127). Bad Heilbrunn: Klinkhardt.
- Leonhard, T. & Rihm, T. (2011). Erhöhung der Reflexionskompetenz durch Begleitveranstaltungen zum Schulpraktikum? Konzeption und Ergebnisse eines Pilotprojekts mit Lehramtsstudierenden. *Lehrerbildung auf dem Prüfstand*, 4(2), 240–270.
- Leuders, T. & Holzäpfel, L. (2011). Kognitive Aktivierung im Mathematikunterricht. *Unterrichtswissenschaft*, 39(3), 213–230.
- Lewis, C., Perry, R. & Hurd, J. (2004). A deeper look at lesson study. *Educational leadership*, 61(5), 18–22.
- Liebsch, K. (Hrsg.). (2010). *Reflexion und Intervention. Zur Theorie und Praxis Schulpraktischer Studien*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Beltz.
- Lipowsky, F. (2006). Auf den Lehrer kommt es an. *Zeitschrift für Pädagogik*, 51, 47–65.
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 69–105). Wiesbaden: Springer.
- London, M. (2003). *Job feedback: Giving, seeking, and using feedback for performance improvement*. London: Psychology Press.
- Lotz, M. & Lipowsky, F. (2015). Die Hattie-Studie und ihre Bedeutung für den Unterricht–Ein Blick auf ausgewählte Aspekte der Lehrer-Schüler-Interaktion. In G. Mehlhorn, F. Schulz & K. Schöppe (Hrsg.), *Begabungen entwickeln & Kreativität fördern* (S. 97–136). München: kopaed.
- Loucks-Horsley, S. & Matsumoto, C. (1999). Research on professional development for teachers of mathematics and science: The state of the scene. *School science and mathematics*, 99(5), 258–271.
- Loughran, J. (2002). Effective reflective practice: In search of meaning in learning about teaching. *Journal of teacher education*, 53(1), 33–43.

- Lüders, M. (2012). „Pädagogisches Unterrichtswissen“ – eine Testkritik. *Zeitschrift für Erziehungswissenschaft*, 15(4), 775–791.
- Luesebrink, I. & Grimminger, E. (2014). Fallorientierte Lehrer- und Lehrerinnenausbildung evaluieren – Überlegungen zur Modellierung reflexiver Kompetenz. In I. Pieper, P. Frei, K. Hausenschild & B. Schmidt-Thieme (Hrsg.), *Was der Fall ist. Beiträge zur Fallarbeit in Bildungsforschung, Lehramtsstudium, Beruf und Ausbildung* (S. 201–212). Wiesbaden: Springer.
- Luft, J. & Ingham, H. (1955). *The Johari window, a graphic model of interpersonal awareness. Proceedings of the western training laboratory in group development*. University of California, Los Angeles, Los Angeles.
- Luhmann, N. (1981). Handlungstheorie und Systemtheorie. In N. Luhmann (Hrsg.), *Soziologische Aufklärung 3. Soziales System, Gesellschaft, Organisation* (S. 50–66). Wiesbaden: Springer.
- Lunetta, V. N., Hofstein, A. & Clough, M. P. (2007). Learning and teaching in the school science laboratory: An analysis of research, theory, and practice. *Handbook of research on science education*, 2, 393–441.
- Magnusson, S., Krajcik, J. & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Hrsg.), *Examining pedagogical content knowledge* (S. 95–132). Wiesbaden: Springer.
- Maier, U. (2010). Formative Assessment – Ein erfolgversprechendes Konzept zur Reform von Unterricht und Leistungsmessung? *Zeitschrift für Erziehungswissenschaft*, 13(2), 293–308.
- Marcos, J.-J. M., Sanchez, E. & Tillema, H. H. (2011). Promoting teacher reflection: What is said to be done. *Journal of Education for teaching*, 37(1), 21–36.
- Maslach, C. & Leiter, M. P. (1999). Teacher Burnout: A Research Agenda. In R. Vandenberg & A. Huberman (Hrsg.), *Understanding and Preventing Teacher Burnout* (295–202). Cambridge: Cambridge University Press.
- Mayer-Frühwirth, G. (2017). Social Video Learning. Kollaboratives Reflektieren in den Schulpraktika. *Forschungsperspektiven*, 9, 69–76.
- Mayr, Eder & Fartacek. (1988). Praxisschock auf Raten. Einphasige Lehrerausbildung und Einstellungsentwicklung. *Unterrichtswissenschaft*, 16(1), 68–82.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Beltz: Weinheim.
- Mc Manus, J. (2011, April). *A new wave of sensation: The critical role emotions play in critical reflection*. Presentation OLKC Making Waves Conference, Kingston Upon Hull, England.
- Meinhardt, C., Rabe, T. & Krey, O. (2015). Quantitative Validierung eines Testinstruments zu Selbstwirksamkeitserwartungen in physikdidaktischen Handlungsfeldern.

- Erste Ergebnisse. In S. Bernholt (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht*. Kiel: IPN.
- Meißner, C., Dohrmann, R. & Nordmeier, V. (2020). Motivationale Wirkungen des Lehrformats Lehr-Lern-Labor. *PhyDid B-Didaktik der Physik-Beiträge zur DPG-Frühjahrstagung*.
- Meyer, H. (2003). Zehn Merkmale guten Unterrichts. *Pädagogik*, 10(2003), 36–43.
- Mezirow, J. How critical reflection triggers transformative learning. In J. Mezirow (Hrsg.), *Fostering Critical Reflection in Adulthood* (S. 1–18). San Francisco: Jossey-Bass.
- Mezirow, J. (1997). *Cognitive processes: contemporary paradigms of learning. Adult learning: A reader*. Sterling, VA: Kogan Page.
- Mezirow, J. (1998). On critical reflection. *Adult education quarterly*, 48(3), 185–198.
- Michel, L. & Conrad, W. (1982). Theoretische Grundlagen psychometrischer Tests. *Grundlagen psychologischer Diagnostik*, 1, 1–129.
- Mietzel, G. (2017). *Pädagogische Psychologie des Lernens und Lehrens*. Göttingen: Hogrefe.
- Mirzaei, F., Phang, F. A. & Kashefi, H. (2014). Measuring Teachers Reflective Thinking Skills. *Procedia-Social and Behavioral Sciences*, 141, 640–647.
- Mönig, S. (2013). *Wissenschaftliche Reflexionsfähigkeit von Lehrenden und Lehramtsstudierenden-theoretische Kompetenzmodellierung und empirische Überprüfung*. Dissertation. Technische Universität Dortmund, Dortmund.
- Moon, J. (2013). *Reflection in learning and professional development: Theory and practice*. London: Routledge.
- Moosbrugger, H. & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion*. Wiesbaden: Springer.
- Mühlhausen. (2012). *Unterrichtsanalyse online. Didaktische Kategorien mit angereicherten Unterrichtsvideos erschließen und überprüfen*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Mühlhausen, U. (2005). Multimediale Unterrichtsdokumente - Reflexion und Analyse von Unterricht. *Journal für LehrerInnenbildung*, (2), 19–25.
- Mühlhausen, U. (2012). Mit der virtuellen Unterrichtshospitation dem Theorie--Praxis-Dilemma der Lehrerbildung entgegenwirken. *Seminar - Lehrerbildung und Schule*, (2), 64–73.
- Munby, H. (1989). Reflection-in-action and reflection-on-action. *Current issues in education*, 9(1), 31–42.
- Münzinger, W. (2001). Lehr--Lern--Labor. Ein Projekt zur Neuorganisation der Lehrerfortbildung im mathematisch-naturwissenschaftlichen Bereich. *Unterricht Physik*, 12(63-64), 72–73.

- Narciss, S. (2006). *Informatives tutorielles Feedback: Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse*. Münster: Waxmann.
- National Research Council. (1996). *National science education standards*. Washington, D. C.: National Academies Press.
- Neber, H. (1998). Kooperatives Lernen. In J. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz.
- Neophytou, L., Koutselini, M. & Kyriakides, L. (2011). The effect of teachers' personal beliefs and emotional intelligence on quality and effectiveness of teaching. In A. Lauriala, R. Rajala, H. Ruokamo & O. Ylitapio-Mäntylä (Hrsg.), *Navigating in Educational Contexts* (S. 207–224). Wiesbaden: Springer.
- Neuhaus, D. (2019). Überlegungen zu einem Reflexionsformat für das Praxissemester Musik. In M. Degeling, N. Franken, S. Freund, S. Greiten & D. Neuhaus (Hrsg.), *Herausforderung Kohärenz: Praxisphasen in der universitären Lehrerbildung. Bildungswissenschaftliche und fachdidaktische Perspektiven*. (S. 222–233). Bad Heilbrunn: Julius Klinkhardt.
- Neuweg, G. H. (2005). Emergenzbedingungen pädagogischer Könnerschaft. In S. Heid & C. Harteis (Hrsg.), *Verwendbarkeit. Ein Qualitätskriterium (erziehungs-)wissenschaftlichen Wissens?* (S. 205–228). Wiesbaden: Springer.
- Neuweg, G. H. (2011). Das Wissen der Wissensvermittler. Problemstellungen, Befunde und Perspektiven der Forschung zum Lehrerwissen. In Terhart, Bennewitz & Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 451–477). Münster: Waxmann.
- Neuweg, G. H. (2014). Das Wissen der Wissensvermittler: Problemstellung, Befunde und Perspektiven der Forschung zum Lehrerwissen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 583–614). Münster: Waxmann.
- Nilsson, P. (2014). When Teaching Makes a Difference: Developing science teachers' pedagogical content knowledge through learning study. *International Journal of Science Education*, 36(11), 1794–1814.
- Nitsche, K. (2014). *UNI-Klassen-Reflexion und Feedback über Unterricht in Videolabors an Schulen*. Dissertation. Ludwig Maximilians Universität, München.
- Nitsche, S., Dickhäuser, O., Fasching, M. S. & Dresel, M. (2011). Rethinking teachers' goal orientations: Conceptual and methodological enhancements. *Learning and Instruction*, 21(4), 574–586.
- Nittel, D. (2011). Von der Profession zur sozialen Welt pädagogisch Tätiger. *Zeitschrift für Pädagogik*, 57, 40–59.
- Noffke, S. E. & Brennan, M. (2005). The Dimensions of Reflection: A Conceptual and Contextual Analysis. *International Journal of Progressive Education*, 1(3), 58–78.

- Nordmeier, Käpnick, Komorek, Leuchter, Neumann, Priemer et al. (2014). *Schülerlabore als Lehr-Lern-Labore: Forschungsorientierte Verknüpfung von Theorie und Praxis in der MINT- Lehrerbildung*: Unveröffentlichter Projektantrag.
- Nowak, A., Ackermann, P. & Borowski, A. (2018). Reflexion von Physikunterricht. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe. Gesellschaft für Didaktik der Chemie und Physik* (S. 838–841). Regensburg: Universität Regensburg.
- Nuthall, G. (2007). *The hidden lives of learners*. Wellington, New Zealand: Nzcer Press Wellington.
- OECD. (2002). *The Measurement of Scientific and Technological Activities Frascati Manual 2002: Proposed Standard Practice for Surveys on Research and Experimental Development*. Cambridge: Cambridge University Press.
- Oelkers, J. (2007). *Kompetenz und Professionalität: Neue Wege in der Lehrerbildung*, Lörrach am Rhein. Zugriff am 05.10.2020.
- Öhlschlegel-Haubrock, Rach & Wolf. (2014). Förderung der Selbstwirksamkeitserwartung in der Studieneingangsphase mittels Mentoring. *Zeitschrift für Hochschulentwicklung ZFHE Jg*, 9(1), 23–35.
- Olander, C. & Olander, M. H. (2013). Professional Development Through the Use of Learning study: contributions to pedagogical content knowledge in biology. *Procedia-Social and Behavioral Sciences*, 89, 205–212.
- Op't Eynde, P., Corte, E. de & Verschaffel, L. Framing students' mathematics-related beliefs. In G. C. Leder, E. Pehkonen & G. Törner (Hrsg.), *Beliefs: A hidden variable in mathematics education?* (S. 13–37). Wiesbaden: Springer.
- Oser. (1992). *Effective and Responsible Teaching: The New Synthesis*. San Francisco: Jossey-Bass.
- Oser. (1997). Standards in der Lehrerbildung. Teil 1: Berufliche Kompetenzen, die hohen Qualitätsmerkmalen entsprechen. *Beiträge zur Lehrerinnen-und Lehrerbildung*, 15(1), 26–37.
- Oser, Zutavern & Patry (1990). Professionelle Lehrermoral: Das „Gelebte Wertesystem“ von LehrerInnen und seine Veränderbarkeit. In L.-M. Alisch, J. Baumert & K. Beck (Hrsg.), *Professionswissen und Professionalisierung* (S. 227–252). Braunschweig: Technische Univ. Seminar für Soziologie und Sozialarbeitswissenschaften, Abteilung Sozialarbeit.
- Osterman, K. F. & Kottkamp, R. B. (2004). *Reflective practice for educators. Professional development to improve student learning*. Thousand Oaks, Calif.: Corwin Press.
- Pang & Lo Ling, M. (2012). Learning study: Helping teachers to use theory, develop professionally, and produce new knowledge to be shared. *Instructional Science*, 40(3), 589–606.

- Pang & Marton. (2003). Beyond "lesson study": Comparing two ways of facilitating the grasp of some economic concepts. *Instructional Science*, 31(3), 175–194.
- Pang & Marton, F. (2005). Learning theory as teaching resource: Enhancing students' understanding of economic concepts. *Instructional Science*, 33(2), 159–191.
- Paris, S. G., Lipson, M. Y. & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8(3), 293–316.
- Park, S. & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in science Education*, 38(3), 261–284.
- Pauli, C. & Reusser, K. (2006). Von international vergleichenden Video Surveys zur videobasierten Unterrichtsforschung und-entwicklung. *Zeitschrift für Pädagogik*, 52(6), 774–798.
- Pauli, C., Reusser, K., Waldis, M. & Grob, U. (2003). Erweiterte Lehr-und Lernformen" im Mathematikunterricht der Deutschschweiz. *Unterrichtswissenschaft*, 31(4), 291–320.
- Pekrun, R. (2005). Research area 3: Students' egagement in science. In OECD (Hrsg.), *Contextual framework for PISA 2006* (S. 41–56).
- Pfadenhauer, M. (2005). *Professionelles Handeln*. Wiesbaden: Verlag für Sozialwissenschaften.
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D. & Neubrand, M. (Hrsg.). (2004). *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- Raab-Steiner, E. & Benesch, M. (2010). *Der Fragebogen: von der Forschungsidee zur SPSS/PASW-Auswertung*. Wien: Facultas.
- Rabe, T., Meinhardt, C. & Krey, O. (2012). Entwicklung eines Instruments zur Erhebung von Selbstwirksamkeitserwartungen in physikdidaktischen Handlungsfeldern. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 293–315.
- Radtke, F.-O. & Webers, H.-E. (1998). Schulpraktische Studien und Zentren für Lehramtsausbildung. Eine Lösung sucht ihr Problem. *Die Deutsche Schule*, 90(1), 199–216.
- Rakoczy, K., Klieme, E., Drollinger-Vetter, B., Lipowsky, F., Pauli, C. & Reusser, K. (2007). Structure as a quality feature in mathematics instruction: Cognitive and motivational effects of a structured organisation of the learning environment vs. a structured presentation of learning content. *Studies on the educational quality of schools. The final report on the DFG Priority Programme*, 101–120.
- Reh, S. (2004). Abschied von der Profession, von Professionalität oder vom Professionellen? Theorien und Forschungen zur Lehrerprofessionalität. *Zeitschrift für Pädagogik*, 50(3), 358–372.

- Rehfeldt, Klempin & Nordmeier (2018). Ergebnisse fächerübergreifender Praxisrelevanz und Reflexionskompetenz in Lehr-Lern-Laboren. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe. Gesellschaft für Didaktik der Chemie und Physik* (S. 544–547). Regensburg: Universität Regensburg.
- Reich, K. (2006). *Konstruktivistische Didaktik: Lehr-und Studienbuch mit Methodenpool*. Weinheim: Beltz.
- Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning sciences*, 13(3), 273–304.
- Reusser, K. (2005). Situiertes Lernen mit Unterrichtsvideos. *journal für lehrerinnen- und lehrerbildung*, 2, 8–18.
- Reusser, K., Pauli, C. & Waldis, M. (2010). *Unterrichtsgestaltung und Unterrichtsqualität. Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht*. Münster: Waxmann Verlag.
- Riebesecker, T. (2018). *Anpassung eines Kodiermanuals zur Messung der Entwicklung der Reflexionskompetenz von Chemielehramtsstudierenden*. Zulassungsarbeit. Universität Regensburg, Regensburg.
- Riese, J. (2009). *Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften*. Berlin: Logos Verlag.
- Riese, J. & Reinhold, P. (2008). Entwicklung und Validierung eines Instruments zur Messung professioneller Handlungskompetenz bei (angehenden) Physiklehrkräften. *Lehrerbildung auf dem Prüfstand*, 1(2), 625–640.
- Riese, J. & Reinhold, P. (2010). Empirische Erkenntnisse zur Struktur professioneller Handlungskompetenz von angehenden Physiklehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 167–187.
- Rivera, R. (2017). The reflective writing continuum: Re-conceptualizing Hatton & Smith's types of reflective writing. *International Journal of Research Studies in Education*, 6(2), 49–67.
- Rodgers, C. (2002). Defining reflection: Another look at John Dewey and reflective thinking. *Teachers college record*, 104(4), 842–866.
- Rost, D. H. (2013). *Interpretation und Bewertung pädagogisch-psychologischer Studien: Eine Einführung*. Bad Heilbrunn: Klinkhardt.
- Roters, B. (2012). *Professionalisierung durch Reflexion in der Lehrerbildung*. Münster: Waxmann.
- Rovegno, I. (1993). Content-knowledge acquisition during undergraduate teacher education: Overcoming cultural templates and learning through practice. *American Educational Research Journal*, 30(3), 611–642.
- Rumann, S. (2005). *Kooperatives experimentieren im Chemieunterricht. Entwicklung und Evaluation einer Interventionsstudie zur Säure-Base-Thematik*. Berlin: Logos Verlag.

- Runesson, U. (2008). Learning to design for learning: The potential of learning study to enhance teachers' and students' learning. In D. Portri & O. Chapman (Hrsg.), *The International Handbook of Mathematics Teacher Education* (S. 153–172). Boston: Brill Sense.
- Runesson, U. & Gustafsson, G. (2012). Sharing and developing knowledge products from Learning Study. *International Journal for Lesson and Learning Studies*, 1(3), 245–260.
- Santagata, R. & Guarino, J. (2011). Using video to teach future teachers to learn from teaching. *Zdm*, 43(1), 133–145.
- Schaarschmidt, U. & Fischer, A. W. (2003). *AVEM–Arbeitsbezogenes Erlebens- und Verhaltensmuster*: Swets Test Services, Frankfurt/M.
- Schädlich, B. (2015). Fachpraktika im Master of Education Französisch aus der Perspektive der Studierenden: Ein Beitrag zur Entwicklung reflexiver Handlungskompetenz? *Zeitschrift für Fremdsprachenforschung*, 26(2), 255–285.
- Schmelzing, S., Wüsten, S., Sandmann, A. & Neuhaus, B. J. (2010). Fachdidaktisches Wissen und Reflektieren im Querschnitt der Biologielehrerbildung. *Zeitschrift für Didaktik der Naturwissenschaften*, (16), S. 189–207.
- Schmidt, M. (2014). *Professionswissen von Sachunterrichtslehrkräften. Zusammenhangsanalyse zur Wirkung von Ausbildungshintergrundshintergrund und Unterrichtserfahrung auf das fachspezifische Professionswissen im Unterrichtsinhalt "Verbrennung"*. Berlin: Logos Verlag.
- Schmitt, A. K. (2016). *Entwicklung und Evaluation einer Chemielehrerfortbildung zum Kompetenzbereich Erkenntnisgewinnung*. Logos Verlag Berlin.
- Schneider, R. M. (2019). Illustrating and Developing Science Teachers' Pedagogical Content Knowledge Through Learning Study. In A. Hume, R. Cooper & A. Borowski (Hrsg.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge für Teaching Science* (S. 165–184). Wiesbaden: Springer.
- Schön, D. A. (1984). *The Reflective Practitioner. How Professionals Think in Action*. New York: Basic Books.
- Schreiber, N., Theyßen, H. & Schecker, H. (2009). Experimentelle Kompetenz messen? *PhyDid A-Physik und Didaktik in Schule und Hochschule*, 3(8), 92–101.
- Schulz, A. (2011). *Experimentierspezifische Qualitätsmerkmale im Chemieunterricht: Eine Videostudie*. Berlin: Logos Verlag Berlin.
- Schüpbach, J. (2007). *Über das Unterrichten reden: die Unterrichtsnachbesprechung in den Lehrpraktika-eine "Nahtstelle von Theorie und Praxis"?* Bern: Haupt Verlag AG.
- Schwendenwein, W. (1990). Professionswissen und Professionalisierung. In L.-M. Aalisch, J. Baumert & K. Beck (Hrsg.), *Professionswissen und Professionalisierung* (Braunschweiger Studien zur Erziehungs- und Sozialarbeitswissenschaft, Bd. 28,

- S. 359–381). Braunschweig: Technische Univ. Seminar für Soziologie und Sozialarbeitswiss. Abt. Sozialarbeitswiss.
- Schwindt, K. (2008). *Lehrpersonen betrachten Unterricht*. Münster: Waxmann Verlag.
- Seago, N. (2003). Using video as an object of inquiry for mathematics teaching and learning. In Brophy (Hrsg.), *Using video in teacher education. (Advances in Research on Teaching)* (S. 259–286). Bingley: Emerald Group Publishing Limited.
- Seidel, T., Prenzel, M., Rimmele, R., Dalehefte, I. M., Herweg, C., Kobarg, M. et al. (2006). Blicke auf den Physikunterricht. Ergebnisse der IPN Videostudie. *Zeitschrift für Pädagogik*, 52(6), 799–821.
- Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M. & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and teacher education*, 27(2), 259–267.
- Seifert, A., Hilligus, A. H. & Schaper, N. (2009). Entwicklung und psychometrische Überprüfung eines Messinstruments zur Erfassung pädagogischer Kompetenzen in der universitären Lehrerbildung. *Lehrerbildung auf dem Prüfstand*, 2(1), 82–103.
- Seifert, A. & Schaper, N. (2012). Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerausbildung. In J. König & A. Seifert (Hrsg.), *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerausbildung* (S. 183–214). Münster: Waxmann Verlag.
- Sherin, M. G. (2007). The development of teachers' professional vision in video clubs. In Goldman, Roy, B. Barron & Sharon (Hrsg.), *Video Research in the Learning Sciences* (S. 383–395). London: Routledge.
- Shiel, C. & Jones, D. (2003). *Reflective Learning and Assessment. A Systematic Study of Reflexive Learning as Evidenced in Student Learning Journals*. Bournemouth University, Bournemouth.
- Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, 15(2), 4–14.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard educational review*, 57(1), 1–23.
- Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153–189.
- Singer, S. R., Hilton, M. I. & Schweingruber, H. A. (2006). *America's Lab Report. Investigation in High School Science*. Washington, D. C.: The National Academies Press.

- Slavin, R. E. (1989). Cooperative learning and student achievement: Six theoretical perspectives. In C. Ames & M. Maehr (Hrsg.), *Advances in Motivation and Achievement: Motivation Enhancing Environments (Advances in Motivation & Achievement)* (S. 161–177). Stamford: JAI Press Inc.
- Slavin, R. E. (1994). Quality, appropriateness, incentive, and time: A model of instructional effectiveness. *International Journal of Educational Research*, 21(2), 141–157.
- Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21(1), 43–69.
- Smith, F. (1995). Let's declare education a disaster and get on with our lives. *Phi Delta Kappan*, 76(8), 584.
- Sparks-Langer, G. M., Simmons, J. M., Pasch, M., Colton, A. & Starko, A. (1990). Reflective pedagogical thinking: How can we promote it and measure it? *Journal of teacher education*, 41(5), 23–32.
- Staub, F. C. & Stern, E. (2002). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. *Journal of educational psychology*, 94(2), 344.
- Steffens, U. & Höfer, D. (2015). Feedback als Instrument des Lernens im Lichte der Hattie-Studie. In C. G. Buhren (Hrsg.), *Handbuch Feedback in der Schule*. Weinheim und Basel: Beltz.
- Stichweh, R. (1992). Professionalisierung, Ausdifferenzierung von Funktionssystemen, Inklusion. In Dewe, Ferchhoff & Radtke (Hrsg.), *Erziehen als Profession. Zur Logik professionellen Handelns in pädagogischen Feldern* (S. 36–48). Wiesbaden: Springer.
- Straßer, P. (2008). *Können erkennen-reflexives Lehren und Lernen in der beruflichen Benachteiligtenförderung: Entwicklung, Erprobung und Evaluation eines reflexiven Lehr-Lerntrainings*. München: W. Bertelsmann Verlag.
- Stroebe, W., Hewstone, M. & Stephenson, G. M. (1996). *Sozialpsychologie*. Berlin: Springer.
- Stürmer, K., Könings, K. D. & Seidel, T. (2013). Declarative knowledge and professional vision in teacher education: Effect of courses in teaching and learning. *British Journal of Educational Psychology*, 83(3), 467–483.
- Sumfleth, E., Rumann, S. & Nicolai, N. (2004). Kooperatives Arbeiten im Chemieunterricht–Gemeinsames Arbeiten in kleinen Gruppen und mit Eltern. *Essener Unikate*, (24), 74–85.
- Sumfleth, E., Wirth, J. & Thillmann, H. (2013). Selbstreguliertes Experimentieren. *nwu-essen*, 10, 163–208.
- Swanson, H. L., O'Connor, J. E. & Cooney, J. B. (1990). An information processing analysis of expert and novice teachers' problem solving. *American Educational Research Journal*, 27(3), 533–556.

- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4(4), 295–312.
- Szogs, Krüger & Korneck (2019). Veränderung der Unterrichtsqualität durch kollegiale Reflexion. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe* (S. 329–332). Regensburg: Universität Regensburg.
- Tan, Y. S. M. & Nashon, S. M. (2013). Promoting teacher learning through learning study discourse: The case of science teachers in Singapore. *Journal of Science Teacher Education*, 24(5), 859–877.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S. et al. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7–28.
- Terhart, E. (2005). Standards für die Lehrerbildung - Ein Kommentar. *Zeitschrift für Pädagogik*, 51, 275–279.
- Terhart, E. (2007). Erfassung und Beurteilung der beruflichen Kompetenz von Lehrkräften. In M. Lüders & J. Wissinger (Hrsg.), *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation* (S. 37–62). Münster: Waxmann.
- Terhart, E. (2009). *Didaktik: Eine Einführung*. Ditzingen: Reclam.
- Terhart, E. (2011). Lehrerberuf und Professionalität: Gewandeltes Begriffsverständnis – neue Herausforderungen. *Zeitschrift für Pädagogik*, 57(57), 202–224.
- Terhart, E., Bennewitz, H. & Rothland, M. (Hrsg.). (2014). *Handbuch der Forschung zum Lehrerberuf*. Münster: Waxmann.
- Thielsch, M. T., Lenzner, T. & Melles, T. (2012). Wie gestalte ich gute Items und Interviewfragen. In M. T. Thielsch & T. Brandenburg (Hrsg.), *Praxis der Wirtschaftspsychologie II. Themen und Fallbeispiele für Studium und Anwendung* (S. 221–240). Münster: MV-Wissenschaft.
- Tobin & Fraser. (2003). *International handbook of science education*. London: Kluwer Academic.
- Tom, A. R. (1985). Inquiring into inquiry-oriented teacher education. *Journal of teacher education*, 36(5), 35–44.
- Tuma, R., Schnettler, B. & Knoblauch, H. (2013). Videographie. In R. Tuma, B. Schnettler & H. Knoblauch (Hrsg.), *Videographie. Einführung in die interpretative Videoanalyse sozialer Situationen* (S. 63–84). Wiesbaden: Springer.
- Ulich, D., Hausser, K., Mayring, P., Strehmel, P., Kandler, M. & Degenhardt, B. (1985). *Psychologie der Krisenbewältigung*. Weinheim: Beltz.
- Valli, L. (1997). Listening to other voices: A description of teacher reflection in the United States. *Peabody journal of Education*, 72(1), 67–88.

- Van Driel, J. H., Verloop, N. & Vos, W. de. (1998). Developing science teachers' pedagogical content knowledge. *Journal of research in science teaching*, 35(6), 673–695.
- Van Manen, M. (1977). Linking ways of knowing with ways of being practical. *Curriculum inquiry*, 6(3), 205–228.
- Van Manen, M. (1991). *The tact of teaching*: Ontario: The Althouse Press.
- Vogelsang, C. (2014). *Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften. Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz*. Berlin: Logos Verlag Berlin.
- Voss, T., Kunina-Habenicht, O., Hoehne, V. & Kunter, M. (2015). Stichwort Pädagogisches Wissen von Lehrkräften: Empirische Zugänge und Befunde. *Zeitschrift für Erziehungswissenschaft*, 18(2), 187–223.
- Voss, T. & Kunter, M. (2011). Pädagogisch-psychologisches Wissen 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. 193–214). Münster: Waxmann.
- Wahser, I. (2008). *Training von naturwissenschaftlichen Arbeitsweisen zur Unterstützung experimenteller Kleingruppenarbeit im Fach Chemie*. Berlin: Logos Verlag.
- Walpuski, M. (2006). *Optimierung von experimenteller Kleingruppenarbeit durch Strukturierungshilfen und Feedback: eine empirische Studie*. Berlin: Logos Verlag.
- Walpuski, M. (2017). Qualitätsmerkmale im naturwissenschaftlichen Unterricht. In Maurer (Hrsg.), *Qualitätsvoller Chemie- und Physikunterricht - normative und empirische Dimensionen. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Regensburg 2017* (S. 27–32).
- Weber, K. E., Prilop, C. N. & Kleinknecht, M. (2019). Effekte einer online-und video-basierten Feedback-Intervention im Praktikum auf die professionelle Wahrnehmung von Klassenführung und Selbstwirksamkeitserwartung von Lehramtsstudierenden. In N. McElvany, W. Bos, H. Holtappels & A. Ohle-Peters (Hrsg.), *Bedingungen und Effekte von Lehrerbildung, Lehrkraftkompetenzen und Lehrkraft Handeln. Dortmunder Symposium der Empirischen Bildungsforschung* (S. 123–136). Münster: Waxmann.
- Wehner, F. & Weber, N. (2018). Erfassung der Reflexionskompetenz bei Lehramtsstudierenden anhand von Fallvignetten. In S. Miller, B. Holler-Nowitzki & B. Kottmann (Hrsg.), *Profession und Disziplin. Grundschulpädagogik im Diskurs* (S. 270–275). Wiesbaden: Springer.
- Weinert, F. (1999). Psychologische Orientierung in der Pädagogik. In H. Röhrs & H. Scheuerl (Hrsg.), *Richtungsstreit in der Erziehungswissenschaft und pädagogische Verständigung. Wilhelm Flitner zur Vollendung seines 100. Lebensjahres am 20. Aug. 1989 gewidmet*. (S. 203–214). Frankfurt a. Main: Lang.
- Weinert, F. (2001). Vergleichende Leistungsmessung in Schulen. Eine umstrittene Selbstverständlichkeit. In F. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S. 17–32). Weinheim: Beltz.

- Weinert, F. & Helmke, A. (1996). Der gute Lehrer: Person, Funktion oder Fiktion? *Zeitschrift für Pädagogik*, 34, 223–233.
- Weinert, F., Schrader, F.-W. & Helmke, A. (1989). Quality of instruction and achievement outcomes. *International Journal of Educational Research*, 13(8), 895–914.
- Wideen, M., Mayer-Smith, J. & Moon, B. (1998). A critical analysis of the research on learning to teach: Making the case for an ecological perspective on inquiry. *Review of educational research*, 68(2), 130–178.
- Wiechmann, J. (Hrsg.). (2010). *Zwölf Unterrichtsmethoden. Vielfalt für die Praxis*. Weinheim: Beltz.
- Windt, A., Hasenkamp, A., Rau, S., Lenske, G. & Rumann, S. (2016). Teaching Abilities of Prospective Teachers at the Beginning of the Practical Phase of Teacher Education. *Pre-Service teacher education*, 13(13), 2188–2195.
- Windt, A. & Lenske, G. (2015a). Entwicklung der Reflexion von Sachunterricht in der zweiten Phase der Lehrerausbildung. In H. Fischler, H. Giest & K. Michalik (Hrsg.), *Bildung im und durch Sachunterricht* (S. 209–216). Bad Heilbrunn: Klinkhardt.
- Windt, A. & Lenske, G. (2015b). Qualität der Sachunterrichtsreflexion im Vorbereitungsdienst. In C. Maurer (Hrsg.), *Authentizität und Lernen-das Fach in der Fachdidaktik* (S. 284–286). Regensburg: Universität Regensburg.
- Wirth, J., Thillmann, H., Küsting, J., Fischer, H. E. & Leutner, D. (2008). Das Schülerexperiment im naturwissenschaftlichen Unterricht. Bedingungen der Lernförderlichkeit einer verbreiteten Lehrmethode aus instruktionspsychologischer Sicht. *Zeitschrift für Pädagogik*, 54(3), 361–375.
- Wirtz, M. A. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe.
- Wischmann, F. (2015, 16. Juni). *Mentoring im fachbezogenen Schulpraktikum: Analyse von Reflexionsgesprächen*. Dissertation. Universität Bremen, Bremen.
- Witner, S. & Tepner, O. (2011). Entwicklung geschlossener Testaufgaben zur Erhebung des fachdidaktischen Wissens von Chemielehrkräften. *Chimica et ceterae artes rerum naturae didacticae*, 37(104), 113–137.
- Wolff, C. E., van den Bogert, N., Jarodzka, H. & Boshuizen, H. P. A. (2015). Keeping an eye on learning: Differences between expert and novice teachers' representations of classroom management events. *Journal of teacher education*, 66(1), 68–85.
- Wong, D., Pugh, K. & Dewey Ideas Group at Michigan State University. (2001). Learning science: A Deweyan perspective. *Journal of research in science teaching*, 38(3), 317–336.
- Wood, K. & Sithamparam, S. (2014). *Realising learning: Teachers' professional development through Lesson and Learning Study*. London: Routledge.

- Wubbels, T. & Korthagen, F. (1990). The effects of a pre-service teacher education program for the preparation of reflective teachers. *Journal of Education for teaching*, 16(1), 29–43.
- Wüsten, S. (2010). *Allgemeine und fachspezifische Merkmale der Unterrichtsqualität im Fach Biologie: Eine Video-und Interventionsstudie*. Berlin: Logos Verlag Berlin.
- Wyss, C. (2013). *Unterricht und Reflexion. Eine mehrperspektivische Untersuchung der Unterrichts- und Reflexionskompetenz von Lehrkräften*. Münster, New York, München, Berlin: Waxmann.
- Yost, D. S., Sentner, S. M. & Forlenza-Bailey, A. (2000). An examination of the construct of critical reflection: Implications for teacher education programming in the 21st century. *Journal of teacher education*, 51(1), 39–49.
- Zeichner, K. & Liston, D. (1985). Varieties of discourse in supervisory conferences. *Teaching and teacher education*, 1(2), 155–174.
- Zeichner, K. & Liston, D. (1996). *Reflective Teaching: An Introduction (Reflective Teaching and the Social Conditions of Schooling Series)*: New York, NY: Routledge Publishing.
- Zierer, K., Busse, V., Wernke, S. & Otterspeer, L. (2015). Feedback in der Schule - Forschungsergebnisse. In C. G. Buhren (Hrsg.), *Handbuch Feedback in der Schule* (S. 31–50). Weinheim und Basel: Beltz.
- Zimmermann, M. & Welzel, M. (2008). Reflexionskompetenz - ein Schlüssel zur naturwissenschaftlichen Frühförderkompetenz (NFFK). *Perspektiven ur pädagogischen Professionalisierung: Aspekte zur Elementarbildung II*, 37(74), 29–36.

10. Anhang

10.1. Verzeichnisse

10.1.1. Abbildungsverzeichnis

Abbildung 1: Drei Basisdimensionen der Unterrichtsqualität (Lotz & Lipowsky, 2015, S. 105).....	4
Abbildung 2: Grunddimensionen der Unterrichtsqualität und deren vermutliche Wirkung (Klieme & Rakoczy, 2008, S. 228)	4
Abbildung 3: Modell für Unterrichtsqualität in den Naturwissenschaften mit exemplarischen Messskalen (H. Fischer & Sumfleth, 2013, IX)	5
Abbildung 4: Angebots-Nutzungs-Modell (Helmke, 2012, S. 71).....	6
Abbildung 5: Das Entscheidungsfeld der Unterrichtsmethoden (Wiechmann, 2010, S. 19)	12
Abbildung 6: Learning cycle (Kolb, 2014, S. 21)	23
Abbildung 7: Reflexionsmodell nach Mezirow, 1995 (Abbildung übernommen von Kitchenham, 2008, S. 115)	25
Abbildung 8: ALACT-Modell (Korthagen et al., 2002b, S. 49)	27
Abbildung 9: Der Kreislauf von Aktion und Reflexion der Aktionsforschung (Altrichter, Posch & Spann, 2018, S. 14).....	29
Abbildung 10: Hierarchisches Stufenmodell der Handlungskompetenz (Frey & Jung, 2011, S. 552).....	34
Abbildung 11: Modell professioneller Handlungskompetenz (Baumert & Kunter, 2006, S. 482).....	37
Abbildung 12: Modell zur Konzeption von Items zur Erfassung des Professionswissens (Tepner et al., 2012, S. 19)	44
Abbildung 13: Das Refined Consensus Model (RCM) of PCK (Carlson & Daehler, 2019, S. 83).....	48
Abbildung 14: Enacted PCK (Carlson & Daehler, 2019, S. 84)	48
Abbildung 15: Basisdimensionen wissenschaftlicher Reflexionsfähigkeit (Mönig, 2013, S. 123).....	53
Abbildung 16: Das ALACT-Modell und die nötige Lehrerbildungskompetenz (Korthagen & Kessels, 1999, S. 13)	54
Abbildung 17: Zyklisch forschendes Lernen im Lehr-Lern-Labor (idealisiert) (Nordmeier et al., 2014)	55
Abbildung 18: Selbstreflexion anhand eigener Videoaufzeichnungen (Nitsche, 2014, S. 87).....	57
Abbildung 19: Ablauf einer Learning Study (Nilsson, 2014, S. 1799)	60
Abbildung 20: Übersicht über die Determinanten des Informationswerts von Feedback (Narciss, 2006, S.81)	64
Abbildung 21: A model of feedback to enhance learning (Hattie & Timperley, 2007, S. 87).....	65
Abbildung 22: Ausgewählte Ergebnisse der Studie (Windt et al., 2016, S. 2193)	81
Abbildung 23: Modell zur Einschätzung von Selbst- und Fremdrelexionen von Fachunterricht (Nowak et al., 2018, S. 221).....	86
Abbildung 24: Leitfragen zur Selbstreflexion (Nowak et al., 2018, S. 223).....	87
Abbildung 25: Ablauf und Durchführung des Hauptseminars im Überblick.....	98

Abbildung 26: Unterrichtsstunde im Lehr-Lern-Labor der Universität Regensburg...	100
Abbildung 27: Analyse der Unterrichtsstunden mittels Videos im Lehr-Lern-Labor der Universität Regensburg	101
Abbildung 28: Position der Kameras und Mikrofone im Lehr-Lern-Labor.....	102
Abbildung 29: Beispielitem aus dem Test zum fachdidaktischen Wissen – Studierende	104
Abbildung 30: Beispielitem aus dem Test zum fachdidaktischen Wissen – Studierende	104
Abbildung 31: Beispielitem aus dem Test zum Fachwissen – Studierende.....	105
Abbildung 32: Beispielitem des Reflexionskompetenztests – konditionales Wissen ..	108
Abbildung 33: Beispielitem des Reflexionskompetenztests – deklaratives Wissen	109
Abbildung 34: Ablaufmodell strukturierender Inhaltsanalyse (allgemein) (Mayring, 2010, S. 93).....	112
Abbildung 35: Beispielitem aus dem Fachwissenstest zum Thema Trennung von Stoffgemischen – Schülerinnen und Schüler.....	114
Abbildung 36: Beispielitem aus dem Fachwissenstest zum Thema Chromatografie – Schülerinnen und Schüler.....	114
Abbildung 37: Inhaltsanalytische Gütekriterien nach Krippendorff (1980, S. 158)	119
Abbildung 38: Formel zur Berechnung der Intra-/Intercoderreliabilität nach Holsti (Holsti, 1969, S. 140)	121
Abbildung 39: Mittlere Scores des Tests zum fachdidaktischen Wissen über drei Messzeitpunkte – Pilotstudie gesamt.....	129
Abbildung 40: Mittlere Scores des Tests zum fachdidaktischen Wissen über drei Messzeitpunkte – Pilotstudie aufgeteilt in Gruppen.....	130
Abbildung 41: Mittlere Scores des Tests zum Fachwissen über drei Messzeitpunkte – Pilotstudie gesamt.....	131
Abbildung 42: Mittlere Scores des Tests zum Fachwissen über drei Messzeitpunkte – Pilotstudie aufgeteilt in Gruppen.....	132
Abbildung 43: Mittlere Scores des Reflexionskompetenztests über drei Messzeitpunkte – Pilotstudie gesamt.....	133
Abbildung 44: Mittlere Scores zum Reflexionskompetenztest über drei Messzeitpunkte – Pilotstudie aufgeteilt in Gruppen.....	133
Abbildung 45: Gesamtpunktzahl Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z	148
Abbildung 46: Gesamtpunktzahl Kodiermanual Pilotstudie – Punkte im Gruppenvergleich – Messzeitpunkte x,y,z	148
Abbildung 47: Reflexionsbreite Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z	149
Abbildung 48: Reflexionsbreite Kodiermanual – Punkte im Gruppenvergleich – Messzeitpunkte x,y,z	150
Abbildung 49: Reflexionstiefe Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z	151
Abbildung 50: Reflexionstiefe Kodiermanual – Punkte im Gruppenvergleich – Messzeitpunkte x,y,z	152
Abbildung 51: Mittlere Scores des Tests zum Fachwissen über drei Messzeitpunkte – SuS Pilotstudie gesamt.....	153
Abbildung 52: Mittlere Scores der Wissenstests über drei Messzeitpunkte – SuS Feedbackgruppe.....	154

10. Anhang

Abbildung 53: Mittlere Scores der Wissenstests über drei Messzeitpunkte – SuS Selbstreflexionsgruppe	155
Abbildung 54: Mittlere Scores des Reflexionskompetenztests über drei Messzeitpunkte – Hauptstudie – Gesamtstichprobe	160
Abbildung 55: Mittlere Scores des Reflexionskompetenztests über drei Messzeitpunkte – Hauptstudie – Gruppenvergleich	161
Abbildung 56: Hauptstudie – Kodiermanual – Mittlere Scores – Gesamtpunktzahl – Gesamtstichprobe – Messzeitpunkte x,y,z	162
Abbildung 57: Hauptstudie – Kodiermanual – Mittlere Scores – Gesamtpunktzahl – Gruppenvergleich – Messzeitpunkte x,y,z	163
Abbildung 58: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Gesamtstichprobe – Messzeitpunkte x,y,z	165
Abbildung 59: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Aufteilung in Gruppen – Messzeitpunkte x,y,z	166
Abbildung 60: Hauptstudie – Mittlere Scores der Kategorien der Reflexionstiefe über drei Messzeitpunkte x,y,z	167
Abbildung 61: Hauptstudie – Kodiermanual – Mittlere Scores – Summe der Reflexionstiefe – Gesamtstichprobe – Messzeitpunkte x,y,z	168
Abbildung 62: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionstiefe – Gruppenvergleiche – Messzeitpunkte x,y,z	169
Abbildung 63: Hauptstudie – Mittlere Scores – fachdidaktisches Wissen – Scores Gesamtstichprobe über drei Messzeitpunkte	171
Abbildung 64: Hauptstudie – Mittlere Scores – Fachwissen – Scores Gesamtstichprobe über drei Messzeitpunkte	172
Abbildung 65: Hauptstudie – Mittlere Scores – Studierendenfragebogen – Zielorientierungen – Gesamtstichprobe über drei Messzeitpunkte	174
Abbildung 66: Hauptstudie – Mittlere Scores – Studierendenfragebogen – Fachspezifische Überzeugungen – Gesamtstichprobe über drei Messzeitpunkte	175
Abbildung 67: Hauptstudie – Mittlere Scores – Studierendenfragebogen – Selbstwirksamkeitserwartungen – Gesamtstichprobe über drei Messzeitpunkte	176
Abbildung 68: Hauptstudie – Mittlere Scores über drei MZP – Wissenstest Schülerinnen und Schüler	179
Abbildung 69: Hauptstudie – Residualer Lernzuwachs über drei Messzeitpunkte – Schülerinnen und Schüler	180
Abbildung 70: Hauptstudie – Gruppenvergleiche – Reflexionskompetenztest – Gesamtstichprobe	181
Abbildung 71: Hauptstudie – Gruppenvergleiche residualer Lernzuwachs – Reflexionskompetenz – Kodiermanual – Gesamtpunktzahl, Reflexionsbreite, Reflexionstiefe	182
Abbildung 72: Hauptstudie – Gruppenvergleiche residualer Lernzuwachs – fachdidaktisches Wissen	184
Abbildung 73: Hauptstudie – Gruppenvergleiche residualer Lernzuwachs – Fachwissen	185
Abbildung 74: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Gesamtpunktzahl	186
Abbildung 75: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Planung von Experimenten	187
Abbildung 76: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Durchführung von Experimenten	187

Abbildung 77: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Didaktische Rekonstruktion	188
Abbildung 78: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Umsetzung der didaktischen Rekonstruktion	189
Abbildung 79: Hauptstudie – Gruppenvergleiche – Residualer Lernzuwachs der Unterrichtswirksamkeit SuS – Gesamtstichprobe	190
Abbildung 80: Hauptstudie – Gruppenvergleiche – Residuale Entwicklung – Unterrichtswirksamkeit – Vergleich der Gruppen über die drei Messzeitpunkte	190
Abbildung 81: Hauptstudie – Feedbackgruppe – Residuale Entwicklung der Unterrichtswirksamkeit über drei Messzeitpunkte	191
Abbildung 82: Hauptstudie – Selbstreflexionsgruppe – Residuale Entwicklung der Unterrichtswirksamkeit über drei Messzeitpunkte	192
Abbildung 83: Hauptstudie – Kodiermanual – Mittlere Scores – Perspektivenübernahme – Gesamtstichprobe – Messzeitpunkte x,y,z	194
Abbildung 84: Hauptstudie – Kodiermanual – Mittlere Scores – Perspektivenübernahme – Gruppenaufteilung – Messzeitpunkte x,y,z	195
Abbildung 85: Hauptstudie – Kodiermanual – Literaturbezug – Mittlere Scores – Gesamtstichprobe – Messzeitpunkte x,y,z	197
Abbildung 86: Hauptstudie – Kodiermanual – Literaturbezug – Mittlere Scores – Gruppenaufteilung – Messzeitpunkte x,y,z	198
Abbildung 87: Berechnung des Stichprobenumfangs mithilfe der Software G*Power	306

10.1.2. Tabellenverzeichnis

Tabelle 1: Qualitätsmerkmale im naturwissenschaftlichen Unterricht (Schulz, 2011, S.13-14).....	8
Tabelle 2: Ebenen der Reflexion von Van Manen (1977), zitiert nach Ballard (2006) .	24
Tabelle 3: Zusammenfassung der Modelle zum Konstrukt Reflexion	29
Tabelle 4: Experten-Novizen-Paradigma für den Lehrberuf (Koch-Priewe, 2002, S. 317)	35
Tabelle 5: Ausgewählte Definitionen zum Konstrukt „pädagogisches Wissens“	41
Tabelle 6: Ausgewählte Konzeptualisierungen des Fachwissens.....	43
Tabelle 7: Wichtige Definitionen des RCM nach Carlson & Daehler (2019).....	46
Tabelle 8: Zusammenfassung ausgewählter Konzeptualisierungen des fachdidaktischen Wissens.....	46
Tabelle 9: Modell von Harrington et al. (1996)	52
Tabelle 10: Das Modell des Johari-Fensters.....	62
Tabelle 11: Four types of practical discourse (Zeichner & Liston, 1985, S. 162 ff.).....	68
Tabelle 12: Ebenen der Reflexion von van Manen, zitiert nach Shiel und Jones (2003, S. 6).....	69
Tabelle 13: Types of reflection (Hatton & Smith, 1995, S. 45)	70
Tabelle 14: Fünfstufige Reflexionsskala nach Bain et al. (1999, S. 60)	71
Tabelle 15: Reflexionsbreite nach Bain et al. (1999, S. 59).....	72
Tabelle 16: Stufenmodell nach Kember (Kember et al., 2000, S. 383 ff.).....	73
Tabelle 17: Modell nach Zimmermann und Welzel (2008)	74
Tabelle 18: Stufenmodell zur Einschätzung der Reflexionstiefe (Leonhard & Rihm, 2011, S. 257).....	75
Tabelle 19: Die Ebenen der reflexiven Praxis (Bräuer, 2016, S. 28)	77

10. Anhang

Tabelle 20: Kodiermanual (Wyss, 2013, S. 128)	78
Tabelle 21: Ergebnisse der prozentualen InterCoderübereinstimmung und der InterCoderreliabilität nach Fleiss' Kappa	79
Tabelle 22: Kategoriensystem (Windt & Lenske, 2015a, S. 214).....	80
Tabelle 23: Zusammenfassung und Bezug auf das selbst erarbeitete Kodiermanual.....	81
Tabelle 24: Ergebnisse der Studie von El-Dib (2007).....	83
Tabelle 25: Stichprobengröße – Studierende der Interventionsgruppen (IG) und der Kontrollgruppe (KG) bzw. Schülerinnen und Schüler (SuS).....	96
Tabelle 26: Durchgeführte Tests mit Quellenangaben	99
Tabelle 27: Themen der Unterrichtseinheiten und zugehörige Experimente nach Anthofer (2017, S. 96)	99
Tabelle 28: Originalquellen des Studierendenfragebogens	105
Tabelle 29: Beispielitem Zielorientierung	106
Tabelle 30: Beispielitem Fachspezifische Überzeugung.....	106
Tabelle 31: Beispielitem Selbstwirksamkeitserwartungen.....	107
Tabelle 32: Beispielitem KFT-Test – Schülerinnen und Schüler (Heller & Perleth, 2000, S. 103).....	115
Tabelle 33: Beispielitem Fachinteresse – Schülerinnen und Schüler.....	115
Tabelle 34: Interpretation der Alpha-Werte (George & Mallery, 2002, S. 231).....	117
Tabelle 35: Ergebnisse der Tests nach Anthofer (2017)	117
Tabelle 36: Items des Konstrukts Lernen von Chemie in der Schule	120
Tabelle 37: Items des Konstrukts In meinem Studium strebe ich danach... ..	120
Tabelle 38: Interpretation des Holsti-Koeffizienten.....	121
Tabelle 39: Werte der Signifikanz nach Bühl (2014, S. 177).....	124
Tabelle 40: Interpretation der Effektstärke d.....	125
Tabelle 41: Interpretation der Effektstärke r	125
Tabelle 42: Interpretation der Effektstärke η^2	125
Tabelle 43: Unterscheidung der Korrelationen nach Pearson und Spearman	126
Tabelle 44: Deskriptive Statistik zum fachdidaktischen Wissen über drei Messzeitpunkte – Scores Pilotstudie gesamt	130
Tabelle 45: Deskriptive Statistik zum fachdidaktischen Wissen über drei Messzeitpunkte – Feedbackgruppe – Scores Pilotstudie	130
Tabelle 46: Deskriptive Statistik zum fachdidaktischen Wissen über drei Messzeitpunkte – Selbstreflexionsgruppe – Scores Pilotstudie.....	130
Tabelle 47: Deskriptive Statistik zum fachdidaktischen Wissen über drei Messzeitpunkte – Kontrollgruppe – Scores Pilotstudie.....	131
Tabelle 48: Deskriptive Statistik zum Fachwissen über drei Messzeitpunkte – Scores Pilotstudie gesamt.....	131
Tabelle 49: Deskriptive Statistik zum Fachwissen über drei Messzeitpunkte – Feedbackgruppe – Scores Pilotstudie	132
Tabelle 50: Deskriptive Statistik zum Fachwissen über drei Messzeitpunkte – Selbstreflexionsgruppe – Scores Pilotstudie.....	132
Tabelle 51: Deskriptive Statistik zum Fachwissen über drei Messzeitpunkte – Kontrollgruppe – Scores Pilotstudie.....	132
Tabelle 52: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Pilotstudie gesamt.....	133
Tabelle 53: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Feedbackgruppe – Scores Pilotstudie	134

Tabelle 54: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Selbstreflexionsgruppe – Scores Pilotstudie.....	134
Tabelle 55: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Kontrollgruppe – Scores Pilotstudie.....	134
Tabelle 56: Persönliche Angaben der Studierenden.....	134
Tabelle 57: Deskriptive Statistik zu Zielorientierungen über drei Messzeitpunkte – Studierendenfragebogen – Pilotstudie gesamt.....	135
Tabelle 58: Deskriptive Statistik zu Zielorientierungen über drei Messzeitpunkte – Studierendenfragebogen – Feedbackgruppe.....	135
Tabelle 59: Deskriptive Statistik zu Zielorientierungen über drei Messzeitpunkte – Studierendenfragebogen – Selbstreflexionsgruppe	136
Tabelle 60: Deskriptive Statistik zu Zielorientierungen über drei Messzeitpunkte – Studierendenfragebogen – Kontrollgruppe	136
Tabelle 61: Deskriptive Statistik zur Subskala Lehren von Chemie in der Schule über drei Messzeitpunkte – Fachspezifische Überzeugungen – Studierendenfragebogen – Pilotstudie gesamt.....	136
Tabelle 62: Deskriptive Statistik – Subskala Lehren von Chemie in der Schule über drei Messzeitpunkte – Fachspezifische Überzeugungen – Studierendenfragebogen – Feedbackgruppe.....	137
Tabelle 63: Deskriptive Statistik – Subskala Lehren von Chemie in der Schule über drei Messzeitpunkte – Fachspezifische Überzeugungen – Studierendenfragebogen – Selbstreflexionsgruppe	137
Tabelle 64: Deskriptive Statistik – Subskala Lehren von Chemie in der Schule über drei Messzeitpunkte – Fachspezifische Überzeugungen – Studierendenfragebogen – Kontrollgruppe	137
Tabelle 65: Deskriptive Statistik – Fachspezifische Überzeugungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Pilotstudie gesamt.....	137
Tabelle 66: Deskriptive Statistik – Fachspezifische Überzeugungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Feedbackgruppe.....	138
Tabelle 67: Deskriptive Statistik – Fachspezifische Überzeugungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Selbstreflexionsgruppe	138
Tabelle 68: Deskriptive Statistik – Fachspezifische Überzeugungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Kontrollgruppe	138
Tabelle 69: Reliabilitäten – Selbstwirksamkeitserwartung	138
Tabelle 70: Deskriptive Statistik – Subskala Planung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Pilotstudie gesamt.....	139
Tabelle 71: Deskriptive Statistik – Subskala Planung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Feedbackgruppe.....	139
Tabelle 72: Deskriptive Statistik – Subskala Planung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Selbstreflexionsgruppe	139
Tabelle 73: Deskriptive Statistik – Subskala Planung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Kontrollgruppe	139
Tabelle 74: Deskriptive Statistik – Subskala Durchführung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Pilotstudie gesamt.....	140

10. Anhang

Tabelle 75: Deskriptive Statistik – Subskala Durchführung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Feedbackgruppe.....	140
Tabelle 76: Deskriptive Statistik – Subskala Durchführung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Selbstreflexionsgruppe	140
Tabelle 77: Deskriptive Statistik – Subskala Durchführung von Experimenten über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Kontrollgruppe	140
Tabelle 78: Deskriptive Statistik – Subskala Didaktische Rekonstruktion unterrichtsrelevanter Inhalte über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Pilotstudie gesamt.....	141
Tabelle 79: Deskriptive Statistik – Subskala Didaktische Rekonstruktion unterrichtsrelevanter Inhalte über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Feedbackgruppe.....	141
Tabelle 80: Deskriptive Statistik – Subskala Didaktische Rekonstruktion unterrichtsrelevanter Inhalte über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Selbstreflexionsgruppe	141
Tabelle 81: Deskriptive Statistik – Subskala Didaktische Rekonstruktion unterrichtsrelevanter Inhalte über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Kontrollgruppe	141
Tabelle 82: Deskriptive Statistik – Subskala Didaktische Rekonstruktion im Unterricht über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Pilotstudie gesamt.....	142
Tabelle 83: Deskriptive Statistik – Subskala Didaktische Rekonstruktion im Unterricht über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Feedbackgruppe.....	142
Tabelle 84: Deskriptive Statistik – Subskala Didaktische Rekonstruktion im Unterricht über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Selbstreflexionsgruppe	142
Tabelle 85: Deskriptive Statistik – Subskala Didaktische Rekonstruktion im Unterricht über drei Messzeitpunkte – Selbstwirksamkeitserwartung – Studierendenfragebogen – Kontrollgruppe	142
Tabelle 86: Deskriptive Statistik – Selbstwirksamkeitserwartungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Pilotstudie gesamt.....	143
Tabelle 87: Deskriptive Statistik – Selbstwirksamkeitserwartungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Feedbackgruppe.....	143
Tabelle 88: Deskriptive Statistik – Selbstwirksamkeitserwartungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Selbstreflexionsgruppe	143
Tabelle 89: Deskriptive Statistik – Selbstwirksamkeitserwartungen gesamt über drei Messzeitpunkte – Studierendenfragebogen – Kontrollgruppe	143
Tabelle 90: Berechnungen zur Intracoderreliabilität	144
Tabelle 91: Berechnungen zur InterCoderreliabilität – Pilotstudie gesamt – Kodierer 2	145
Tabelle 92: Berechnungen zur InterCoderreliabilität – Pilotstudie Subskalen – Kodierer 2	145
Tabelle 93: Berechnungen zur InterCoderreliabilität – Hauptstudie gesamt – Kodierer2	146

Tabelle 94: Berechnungen zur InterCoderreliabilität – Hauptstudie Subskalen – Kodierer 2	146
Tabelle 95: Berechnungen zur InterCoderreliabilität – Pilotstudie gesamt – Kodierer 3	146
Tabelle 96: Berechnungen zur InterCoderreliabilität – Pilotstudie Subskalen – Kodierer 3	146
Tabelle 97: Berechnungen zur InterCoderreliabilität – Hauptstudie gesamt – Kodierer 3	147
Tabelle 98: Berechnungen zur InterCoderreliabilität – Hauptstudie Subskalen – Kodierer 3	147
Tabelle 99: Deskriptive Statistik – Gesamtpunktzahl Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z	148
Tabelle 100: Deskriptive Statistik – Gesamtpunktzahl Kodiermanual Pilotstudie – Punkte Feedbackgruppe – Messzeitpunkte x,y,z	148
Tabelle 101: Deskriptive Statistik – Gesamtpunktzahl Kodiermanual Pilotstudie – Punkte Selbstreflexionsgruppe – Messzeitpunkte x,y,z	149
Tabelle 102: Deskriptive Statistik – Gesamtpunktzahl Kodiermanual Pilotstudie – Punkte Kontrollgruppe – Messzeitpunkte x,y,z	149
Tabelle 103: Deskriptive Statistik – Reflexionsbreite Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z	150
Tabelle 104: Deskriptive Statistik – Reflexionsbreite Kodiermanual – Punkte Feedbackgruppe – Messzeitpunkte x,y,z	150
Tabelle 105: Deskriptive Statistik – Reflexionsbreite Kodiermanual – Punkte Selbstreflexionsgruppe – Messzeitpunkte x,y,z	150
Tabelle 106: Deskriptive Statistik – Reflexionsbreite Kodiermanual – Punkte Kontrollgruppe – Messzeitpunkte x,y,z	151
Tabelle 107: Deskriptive Statistik – Reflexionstiefe Kodiermanual – Punkte Pilotstudie gesamt – Messzeitpunkte x,y,z	151
Tabelle 108: Deskriptive Statistik – Reflexionstiefe Kodiermanual – Punkte Feedbackgruppe – Messzeitpunkte x,y,z	152
Tabelle 109: Deskriptive Statistik – Reflexionstiefe Kodiermanual – Punkte Selbstreflexionsgruppe – Messzeitpunkte x,y,z	152
Tabelle 110: Deskriptive Statistik – Reflexionstiefe Kodiermanual – Punkte Kontrollgruppe – Messzeitpunkte x,y,z	152
Tabelle 111: Deskriptive Statistik – Mittelwertvergleiche der Wissenstests über drei Messzeitpunkte – SuS Pilotstudie gesamt	153
Tabelle 112: Deskriptive Statistik – Mittelwertvergleiche der Wissenstests über drei Messzeitpunkte – SuS Feedbackgruppe	154
Tabelle 113: Deskriptive Statistik – Mittelwertvergleiche der Wissenstests über drei Messzeitpunkte – SuS Selbstreflexionsgruppe	155
Tabelle 114: Deskriptive Statistik zum Fachinteresse – Pilotstudie – Schülerinnen und Schüler	156
Tabelle 115: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Hauptstudie – Scores Gesamtstichprobe	160
Tabelle 116: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Hauptstudie – Feedbackgruppe	161
Tabelle 117: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Hauptstudie – Selbstreflexionsgruppe	161

10. Anhang

Tabelle 118: Deskriptive Statistik zum Reflexionskompetenztest über drei Messzeitpunkte – Hauptstudie – Kontrollgruppe	161
Tabelle 119: Ergebnisse des Reflexionskompetenztests – Gruppenvergleiche nichtparametrisch über insgesamt drei Messzeitpunkte – Gesamtstudie	162
Tabelle 120: Hauptstudie – Kodiermanual – Deskriptive Statistik – Gesamtpunktzahl – Gesamtstichprobe – Messzeitpunkte x,y,z	163
Tabelle 121: Hauptstudie – Kodiermanual – Deskriptive Statistik – Gesamtpunktzahl – Feedbackgruppe – Messzeitpunkte x,y,z	163
Tabelle 122: Hauptstudie – Kodiermanual – Deskriptive Statistik – Gesamtpunktzahl – Selbstreflexionsgruppe – Messzeitpunkte x,y,z	164
Tabelle 123: Hauptstudie – Kodiermanual – Deskriptive Statistik – Gesamtpunktzahl – Kontrollgruppe – Messzeitpunkte x,y,z	164
Tabelle 124: Gesamtpunktzahl Kodiermanual – Vergleich mittlere Scores der Messzeitpunkte x,y,z – Aufteilung in Gruppen – nichtparametrisch	164
Tabelle 125: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Gesamtstichprobe – Messzeitpunkte x,y,z	165
Tabelle 126: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Feedbackgruppe – Messzeitpunkte x,y,z	166
Tabelle 127: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Selbstreflexionsgruppe – Messzeitpunkte x,y,z	166
Tabelle 128: Hauptstudie – Kodiermanual – Mittlere Scores – Reflexionsbreite – Kontrollgruppe – Messzeitpunkte x,y,z	166
Tabelle 129: Wissenszuwachs Reflexionsbreite – Aufteilung in Gruppen und Messzeitpunkte x,y,z	167
Tabelle 130: Hauptstudie – Kodiermanual – Deskriptive Statistik – Summe der Reflexionstiefe – Gesamtstichprobe – Messzeitpunkte x,y,z	168
Tabelle 131: Hauptstudie – Kodiermanual – Deskriptive Statistik – Reflexionstiefe – Feedbackgruppe – Messzeitpunkte x,y,z	169
Tabelle 132: Hauptstudie – Kodiermanual – Deskriptive Statistik – Reflexionstiefe – Selbstreflexionsgruppe – Messzeitpunkte x,y,z	169
Tabelle 133: Hauptstudie – Kodiermanual – Deskriptive Statistik – Reflexionstiefe – Kontrollgruppe – Messzeitpunkte x,y,z	169
Tabelle 134: Wissenszuwachs Reflexionstiefe – Aufteilung in Gruppen und Messzeitpunkte x,y,z	170
Tabelle 135: Hauptstudie – Deskriptive Statistik – fachdidaktisches Wissen – Scores Gesamtstichprobe über drei Messzeitpunkte	171
Tabelle 136: Hauptstudie – Deskriptive Statistik – Fachwissen – Scores Gesamtstichprobe über drei Messzeitpunkte	172
Tabelle 137: Hauptstudie – Korrelationen des residualen Lernzuwachses – fachdidaktisches Wissen und Fachwissen mit Kodiermanual – Gesamtstichprobe	173
Tabelle 138: Hauptstudie – Studierendenfragebogen – Reliabilitäten – Gesamtstichprobe	173
Tabelle 139: Hauptstudie – Deskriptive Statistik – Studierendenfragebogen – Zielorientierungen – Gesamtstichprobe über drei Messzeitpunkte	174
Tabelle 140: Hauptstudie – Deskriptive Statistik – Studierendenfragebogen – Fachspezifische Überzeugungen – Gesamtstichprobe über drei Messzeitpunkte	175
Tabelle 141: Hauptstudie – Deskriptive Statistik – Studierendenfragebogen – Selbstwirksamkeitserwartungen – Gesamtstichprobe über drei Messzeitpunkte	177
Tabelle 142: Korrelationen des residualen Lernzuwachses der Gesamtstichprobe	178

Tabelle 143: Hauptstudie – Deskriptive Statistik über drei MZP – Scores Wissenstest Schülerinnen und Schüler.....	179
Tabelle 144: Mittelwertvergleiche - Wissenstest Schülerinnen und Schüler über drei Messzeitpunkte	180
Tabelle 145: Hauptstudie – Deskriptive Statistik – Residualer Lernzuwachs über drei Messzeitpunkte – Schülerinnen und Schüler.....	180
Tabelle 146: Hauptstudie – Gruppenvergleiche – Reflexionskompetenztest – Deskriptive Statistik – Gesamtstichprobe	181
Tabelle 147: Hauptstudie – Gruppenvergleiche – Deskriptive Statistik – Kodiermanual – Gesamtpunktzahl	182
Tabelle 148: Hauptstudie – Gruppenvergleiche – Deskriptive Statistik – Kodiermanual – Reflexionsbreite.....	182
Tabelle 149: Hauptstudie – Gruppenvergleiche – Deskriptive Statistik – Kodiermanual – Reflexionstiefe.....	182
Tabelle 150: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche residualer Lernzuwachs – fachdidaktisches Wissen	184
Tabelle 151: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche residualer Lernzuwachs – Fachwissen	185
Tabelle 152: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Gesamtpunktzahl	186
Tabelle 153: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Planung von Experimenten.....	187
Tabelle 154: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Durchführung von Experimenten	188
Tabelle 155: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Didaktische Rekonstruktion	188
Tabelle 156: Hauptstudie – Deskriptive Statistik – Gruppenvergleiche – Residuale Entwicklung – Selbstwirksamkeit – Umsetzung der didaktischen Rekonstruktion.....	189
Tabelle 157: Hauptstudie – Deskriptive Statistik – Gruppenvergleich – Residualer Lernzuwachs der Unterrichtswirksamkeit SuS – Gesamtstichprobe.....	190
Tabelle 158: Hauptstudie – Deskriptive Statistik – Residualer Lernzuwachs der Unterrichtswirksamkeit – Vergleich der Gruppen über die drei Messzeitpunkte	191
Tabelle 159: Hauptstudie – Feedbackgruppe – Deskriptive Statistik – Residuale Entwicklung der Unterrichtswirksamkeit über drei Messzeitpunkte	192
Tabelle 160: Hauptstudie – Selbstreflexionsgruppe – Deskriptive Statistik – Residuale Entwicklung der Unterrichtswirksamkeit über drei Messzeitpunkte	192
Tabelle 161: Hauptstudie – Korrelationen des residualen Lernzuwachses – fachdidaktisches Wissen und Fachwissen mit Kodiermanual – Ausschluss IG1.....	193
Tabelle 162: Hauptstudie – Kodiermanual – Deskriptive Statistik – Perspektivenübernahme – Gesamtstichprobe – Messzeitpunkte x,y,z	194
Tabelle 163: Hauptstudie – Kodiermanual – Deskriptive Statistik – Perspektivenübernahme – Feedbackgruppe – Messzeitpunkte x,y,z	195
Tabelle 164: Hauptstudie – Kodiermanual – Deskriptive Statistik – Perspektivenübernahme – Selbstreflexionsgruppe – Messzeitpunkte x,y,z	195
Tabelle 165: Hauptstudie – Kodiermanual – Deskriptive Statistik – Perspektivenübernahme – Kontrollgruppe – Messzeitpunkte x,y,z	195
Tabelle 166: Hauptstudie – Kodiermanual – Perspektivenübernahme – Wissenszuwachs – nichtparametrisch – Gruppenaufteilung – Messzeitpunkte x,y,z	196

10. Anhang

Tabelle 167: Hauptstudie – Kodiermanual – Perspektivenübernahme – residualer Lernzuwachs – Gruppenaufteilung.....	196
Tabelle 168: Hauptstudie – Kodiermanual – Literaturbezug – Deskriptive Statistik – Gesamtstichprobe – Messzeitpunkte x,y,z	197
Tabelle 169: Hauptstudie – Kodiermanual – Literaturbezug – Deskriptive Statistik – Feedbackgruppe – Messzeitpunkte x,y,z	198
Tabelle 170: Hauptstudie – Kodiermanual – Literaturbezug – Deskriptive Statistik – Selbstreflexionsgruppe – Messzeitpunkte x,y,z	198
Tabelle 171: Hauptstudie – Kodiermanual – Literaturbezug – Deskriptive Statistik – Kontrollgruppe – Messzeitpunkte x,y,z.....	198
Tabelle 172: Hauptstudie – Kodiermanual – Literaturbezug – Wissenszuwachs – nichtparametrisch – Gruppenaufteilung – Messzeitpunkte x,y,z	199
Tabelle 173: Hauptstudie – Gruppenvergleiche – Deskriptive Statistik – Kodiermanual – Literaturbezug.....	199
Tabelle 174: Berechnungen zur Kriteriumsvalidität des Reflexionskompetenztests ...	306
Tabelle 175: Berechnungen zur korrelativen Gültigkeit des Kodiermanuals.....	306
Tabelle 176: Reliabilitätsanalyse – PCK-Test – Pilot- und Hauptstudie.....	307
Tabelle 177: Reliabilitätsanalyse – CK-Test – Pilotstudie.....	308
Tabelle 178: Reliabilitätsanalyse – CK-Test – Pilot- und Hauptstudie.....	308
Tabelle 179: Reliabilitätsanalyse – Reflexionskompetenztest – Pilotstudie	308
Tabelle 180: Reliabilitätsanalyse – Studierendenfragebogen – Zielorientierungen – Pilotstudie	309
Tabelle 181: Reliabilitätsanalyse – Studierendenfragebogen – Fachspezifische Überzeugungen – Subskala Chemie als Wissenschaft – Pilotstudie	309
Tabelle 182: Reliabilitätsanalyse – Studierendenfragebogen – Fachspezifische Überzeugungen – Subskala Lehrern von Chemie in der Schule – Pilotstudie	309
Tabelle 183: Reliabilitätsanalyse – Studierendenfragebogen – Fachspezifische Überzeugungen – Subskala Lernen von Chemie in der Schule – Pilotstudie	310
Tabelle 184: Reliabilitätsanalyse – Studierendenfragebogen – Fachspezifische Überzeugungen – Gesamt.....	310
Tabelle 185: Reliabilitätsanalyse – Selbstwirksamkeitserwartung – Subskala Planung von Experimenten – Pilotstudie.....	310
Tabelle 186: Reliabilitätsanalyse – Selbstwirksamkeitserwartung – Subskala Durchführung von Experimenten – Pilotstudie	311
Tabelle 187: Reliabilitätsanalyse – Selbstwirksamkeitserwartung – Subskala Didaktische Rekonstruktion unterrichts-relevanter Inhalte – Pilotstudie.....	311
Tabelle 188: Reliabilitätsanalyse – Selbstwirksamkeitserwartung – Subskala Didaktische Rekonstruktion im Unterricht – Pilotstudie.....	311
Tabelle 189: Reliabilitätsanalyse – Selbstwirksamkeitserwartung – Gesamt – Pilotstudie	311
Tabelle 190: Kodiermanual – Intracoderreliabilität – Pilotstudie	313
Tabelle 191: Intercoderreliabilität – Kodiermanual – Pilotstudie gesamt – Arbeitsschritt 2	313
Tabelle 192: Intercoderreliabilität – Kodiermanual – Reflexionsbreite – Pilotstudie..	313
Tabelle 193: Intercoderreliabilität – Kodiermanual – Reflexionstiefe – Pilotstudie....	313
Tabelle 194: Intercoderreliabilität – Kodiermanual – Hauptstudie gesamt.....	313
Tabelle 195: Intercoderreliabilität – Kodiermanual – Reflexionsbreite – Hauptstudie	314
Tabelle 196: Intercoderreliabilität – Kodiermanual – Reflexionstiefe – Hauptstudie..	314

Tabelle 197: Intercorderreliabilität – Kodiermanual – Pilotstudie gesamt – Arbeitsschritt 3	314
Tabelle 198: Intercorderreliabilität – Kodiermanual – Reflexionsbreite – Pilotstudie	314
Tabelle 199: Intercorderreliabilität – Kodiermanual – Reflexionstiefe – Pilotstudie ..	314
Tabelle 200: Intercorderreliabilität – Kodiermanual – Hauptstudie gesamt	314
Tabelle 201: Intercorderreliabilität – Kodiermanual – Reflexionsbreite – Hauptstudie	315
Tabelle 202: Intercorderreliabilität – Kodiermanual – Reflexionstiefe – Hauptstudie	315
Tabelle 203: Reliabilitätsanalyse – CK-Test – SuS	315
Tabelle 204: Reliabilitätsanalyse – Kognitiver Fähigkeitentest – SuS.....	315
Tabelle 205: Kognitiver Fähigkeitentest – Normalverteilung.....	316
Tabelle 206: Kognitiver Fähigkeitentest – Test auf Unterscheidung der Gruppen – Parametrisch	316
Tabelle 207: Kognitiver Fähigkeitentest – Test auf Unterscheidung der Gruppen – Nichtparametrisch.....	316
Tabelle 208: Reliabilitätsanalyse – Fachinteresse – Subskala Freude und Interesse an Chemie.....	316
Tabelle 209: Reliabilitätsanalyse – Fachinteresse – Subskala Chemiebezogenes Fähigkeitsselbstkonzept.....	316
Tabelle 210: Reliabilitätsanalyse – Fachinteresse gesamt.....	317
Tabelle 211: Fachinteresse – Test auf Normalverteilung.....	317
Tabelle 212: Fachinteresse – Test auf Unterscheidung der Gruppen – Parametrisch..	317
Tabelle 213: Überprüfung der Unterscheidung von Pilot- und Hauptstudie – Reflexionskompetenztest.....	317
Tabelle 214: Überprüfung der Unterscheidung von Pilot- und Hauptstudie – Fachwissen	317
Tabelle 215: Überprüfung der Unterscheidung von Pilot- und Hauptstudie – Fachdidaktisches Wissen.....	318
Tabelle 216: Überprüfung der Unterscheidung von Pilot- und Hauptstudie – Fachspezifische Überzeugungen	318
Tabelle 217: Überprüfung der Unterscheidung von Pilot- und Hauptstudie – Selbstwirksamkeitserwartung.....	318
Tabelle 218: Reliabilitätsanalyse – Reflexionskompetenztest – Hauptstudie – 17 Items	318
Tabelle 219: Reliabilitätsanalyse – Reflexionskompetenztest – Hauptstudie – 12 Items	318
Tabelle 220: Reliabilitätsanalyse – Reflexionskompetenztest – Selbstreflexions- und Kontrollgruppe	319
Tabelle 221: Reflexionskompetenztest – Test auf Normalverteilung	319
Tabelle 222: Reflexionskompetenztest – Mittelwertsvergleich – Nichtparametrisch..	319
Tabelle 223: Reflexionskompetenztest – Mittelwertsvergleich – Nichtparametrisch – Feedbackgruppe.....	319
Tabelle 224: Reflexionskompetenztest – Mittelwertsvergleich – Nichtparametrisch – Selbstreflexionsgruppe	319
Tabelle 225: Reflexionskompetenztest – Mittelwertsvergleich – Nichtparametrisch – Kontrollgruppe	320
Tabelle 226: Reflexionskompetenztest – Mittelwertsvergleich – Parametrisch	320
Tabelle 227: Reflexionskompetenztest – Mittelwertsvergleich – Parametrisch – Feedbackgruppe.....	320

10. Anhang

Tabelle 228: Reflexionskompetenztest – Mittelwertsvergleich – Parametrisch – Selbstreflexionsgruppe	320
Tabelle 229: Reflexionskompetenztest – Mittelwertsvergleich – Parametrisch – Kontrollgruppe	320
Tabelle 230: Ergebnisse des Reflexionskompetenztests – Gruppenvergleiche parametrisch – Gesamtstudie.....	320
Tabelle 231: Kodiermanual – Gesamtpunktzahl – Test auf Normalverteilung.....	321
Tabelle 232: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Nichtparametrisch – Gesamtstichprobe.....	321
Tabelle 233: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Parametrisch – Gesamtstichprobe	321
Tabelle 234: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Nichtparametrisch – Feedbackgruppe	321
Tabelle 235: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Nichtparametrisch – Selbstreflexionsgruppe.....	321
Tabelle 236: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Nichtparametrisch – Kontrollgruppe	322
Tabelle 237: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Parametrisch – Feedbackgruppe.....	322
Tabelle 238: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Parametrisch – Selbstreflexionsgruppe	322
Tabelle 239: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Parametrisch – Kontrollgruppe	322
Tabelle 240: Kodiermanual – Reflexionsbreite – Test auf Normalverteilung	322
Tabelle 241: Kodiermanual – Reflexionsbreite – Mittelwertsvergleich – Nichtparametrisch – Gesamtstichprobe.....	322
Tabelle 242: Kodiermanual – Reflexionsbreite – Mittelwertsvergleich – Nichtparametrisch – Feedbackgruppe	323
Tabelle 243: Kodiermanual – Reflexionsbreite – Mittelwertsvergleich – Nichtparametrisch – Selbstreflexionsgruppe.....	323
Tabelle 244: Kodiermanual – Reflexionsbreite – Mittelwertsvergleich – Nichtparametrisch – Kontrollgruppe	323
Tabelle 245: Verteilung der Kategorien der Reflexionstiefe.....	323
Tabelle 246: Kodiermanual – Reflexionstiefe – Test auf Normalverteilung	324
Tabelle 247: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Nichtparametrisch – Gesamtstichprobe	324
Tabelle 248: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Parametrisch – Gesamtstichprobe	324
Tabelle 249: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Nichtparametrisch – Feedbackgruppe.....	324
Tabelle 250: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Parametrisch – Feedbackgruppe.....	324
Tabelle 251: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Nichtparametrisch – Selbstreflexionsgruppe	325
Tabelle 252: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Parametrisch – Selbstreflexionsgruppe	325
Tabelle 253: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Nichtparametrisch – Kontrollgruppe	325

Tabelle 254: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Parametrisch – Kontrollgruppe	325
Tabelle 255: PCK-Test – Test auf Normalverteilung.....	325
Tabelle 256: PCK-Test – Mittelwertsvergleich – Nichtparametrisch – Gesamtstichprobe	326
Tabelle 257: PCK-Test – Mittelwertsvergleich – Parametrisch – Gesamtstichprobe ..	326
Tabelle 258: CK-Test – Test auf Normalverteilung.....	326
Tabelle 259: CK-Test – Mittelwertsvergleich – Nichtparametrisch – Gesamtstichprobe	326
Tabelle 260: CK-Test – Mittelwertsvergleich – Parametrisch – Gesamtstichprobe	326
Tabelle 261: Residualer Lernzuwachs – Test auf Normalverteilung.....	327
Tabelle 262: Parametrische Korrelationen	327
Tabelle 263: Nichtparametrische Korrelationen.....	327
Tabelle 264: Reliabilitäten – Zielorientierungen.....	328
Tabelle 265: Reliabilitäten – Fachspezifische Überzeugungen.....	328
Tabelle 266: Reliabilitäten – Selbstwirksamkeitserwartungen	328
Tabelle 267: Zielorientierungen – Test auf Normalverteilung.....	328
Tabelle 268: Fachspezifische Überzeugungen – Test auf Normalverteilung.....	329
Tabelle 269: Selbstwirksamkeitserwartung – Test auf Normalverteilung	329
Tabelle 270: Zielorientierungen – Mittelwertsvergleich – Nichtparametrischer Test .	329
Tabelle 271: Zielorientierungen – Mittelwertsvergleich – Parametrischer Test	329
Tabelle 272: Fachspezifische Überzeugung – Mittelwertsvergleich – Nichtparametrischer Test	330
Tabelle 273: Fachspezifische Überzeugung – Mittelwertsvergleich – Parametrischer Test	330
Tabelle 274: Selbstwirksamkeitserwartung – Mittelwertsvergleich – Nichtparametrischer Test	330
Tabelle 275: Selbstwirksamkeitserwartung – Mittelwertsvergleich – Parametrischer Test	330
Tabelle 276: Standardisierte Residuen – Test auf Normalverteilung.....	331
Tabelle 277: Korrelationen – Parametrisch	331
Tabelle 278: Korrelationen – Nichtparametrisch	333
Tabelle 279: Wissenstest – SuS – Test auf Normalverteilung MZP1	334
Tabelle 280: Wissenstest – SuS – Test auf Normalverteilung MZP2	334
Tabelle 281: Wissenstest – SuS – Test auf Normalverteilung MZP3	334
Tabelle 282: Wissenstest – SuS – Mittelwertvergleiche – Nichtparametrisch – MZP1	334
Tabelle 283: Wissenstest – SuS – Mittelwertvergleiche – Nichtparametrisch – MZP2	334
Tabelle 284: Wissenstest – SuS – Mittelwertvergleiche – Nichtparametrisch - MZP3	334
Tabelle 285: Wissenstest – SuS – Mittelwertvergleiche – Parametrisch – MZP1	335
Tabelle 286: Wissenstest – SuS – Mittelwertvergleiche – Parametrisch – MZP2	335
Tabelle 287: Wissenstest – SuS – Mittelwertvergleiche – Parametrisch – MZP3	335
Tabelle 288: Residualer Lernzuwachs – SuS – Berechnung von Unterschieden zwischen den drei Messzeitpunkten	335
Tabelle 289: Residualer Lernzuwachs – SuS – Post-Hoc Test	335
Tabelle 290: Reflexionskompetenztest – Residualer Lernzuwachs – Test auf Normalverteilung.....	336

10. Anhang

Tabelle 291: Reflexionskompetenztest – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2	336
Tabelle 292: Reflexionskompetenztest – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG.....	336
Tabelle 293: Kodiermanual – Residualer Lernzuwachs – Test auf Normalverteilung	337
Tabelle 294: Kodiermanual – Gesamtpunktzahl – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. IG2.....	337
Tabelle 295: Kodiermanual – Gesamtpunktzahl – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG	337
Tabelle 296: Kodiermanual – Gesamtpunktzahl – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2	337
Tabelle 297: Kodiermanual – Gesamtpunktzahl – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG.....	338
Tabelle 298: Kodiermanual – Reflexionsbreite – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2	338
Tabelle 299: Kodiermanual – Reflexionsbreite – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG.....	338
Tabelle 300: Kodiermanual – Reflexionstiefe – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. IG2.....	338
Tabelle 301: Kodiermanual – Reflexionstiefe – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG	339
Tabelle 302: Kodiermanual – Reflexionstiefe – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2	339
Tabelle 303: Kodiermanual – Reflexionstiefe – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG.....	339
Tabelle 304: PCK-Test – Residualer Lernzuwachs – Test auf Normalverteilung	339
Tabelle 305: PCK-Test – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. IG2	339
Tabelle 306: PCK-Test – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG	340
Tabelle 307: PCK-Test – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2.....	340
Tabelle 308: PCK-Test – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG	340
Tabelle 309: CK-Test – Residualer Lernzuwachs – Test auf Normalverteilung	340
Tabelle 310: CK-Test – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. IG2	341
Tabelle 311: CK-Test – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG	341
Tabelle 312: CK-Test – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2	341
Tabelle 313: CK-Test – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG	341
Tabelle 314: Selbstwirksamkeitserwartung – Residualer Lernzuwachs – Test auf Normalverteilung.....	342
Tabelle 315: Selbstwirksamkeitserwartung – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2	342
Tabelle 316: Selbstwirksamkeitserwartung – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG.....	342
Tabelle 317: Selbstwirksamkeitserwartung – Planung von Experimenten – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2	342
Tabelle 318: Selbstwirksamkeitserwartung – Planung von Experimenten – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG	343
Tabelle 319: Selbstwirksamkeitserwartung – Durchführung von Experimenten – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2.....	343

Tabelle 320: Selbstwirksamkeitserwartung – Durchführung von Experimenten – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG	343
Tabelle 321: Selbstwirksamkeitserwartung – Didaktische Rekonstruktion – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2	344
Tabelle 322: Selbstwirksamkeitserwartung – Didaktische Rekonstruktion – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG	344
Tabelle 323: Selbstwirksamkeitserwartung – Umsetzung der didaktischen Rekonstruktion – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2	344
Tabelle 324: Selbstwirksamkeitserwartung – Umsetzung der didaktischen Rekonstruktion – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG.....	345
Tabelle 325: Unterrichtsqualität – Residualer Lernzuwachs – Test auf Normalverteilung	345
Tabelle 326: Unterrichtsqualität – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2 – Gesamtstichprobe	345
Tabelle 327: Unterrichtsqualität – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2 – MZP1.....	345
Tabelle 328: Unterrichtsqualität – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2 – MZP2	346
Tabelle 329: Unterrichtsqualität – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2 – MZP3	346
Tabelle 330: Unterrichtsqualität – Residualer Lernzuwachs – Unterschied in den drei MZP – Feedbackgruppe	346
Tabelle 331: Unterrichtsqualität – Residualer Lernzuwachs – Unterschied in den drei MZP – Selbstreflexionsgruppe	346
Tabelle 332: Unterrichtsqualität – Residualer Lernzuwachs – Unterschied in den drei MZP – Post-Hoc-Test – Selbstreflexionsgruppe	347
Tabelle 333: Reflexionskompetenztest – Residualer Lernzuwachs – Vergleich IG2 vs. KG	347
Tabelle 334: Residualer Lernzuwachs – Test auf Normalverteilung – Ausschluss IG1	347
Tabelle 335: Residualer Lernzuwachs – Parametrische Korrelationen – Ausschluss IG1	348
Tabelle 336: Residualer Lernzuwachs – Nichtparametrische Korrelationen – Ausschluss IG1	348
Tabelle 337: Kodiermanual – Perspektivenübernahme – Test auf Normalverteilung .	349
Tabelle 338: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Nichtparametrisch.....	349
Tabelle 339: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Parametrisch	349
Tabelle 340: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Nichtparametrisch – Feedbackgruppe	349
Tabelle 341: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Parametrisch – Feedbackgruppe.....	349
Tabelle 342: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Nichtparametrisch – Selbstreflexionsgruppe	350
Tabelle 343: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Parametrisch – Feedbackgruppe.....	350
Tabelle 344: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Nichtparametrisch – Kontrollgruppe	350

10. Anhang

Tabelle 345: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Parametrisch – Kontrollgruppe	350
Tabelle 346: Perspektivenübernahme – Residualer Lernzuwachs – IG1 vs. IG2	350
Tabelle 347: Perspektivenübernahme – Residualer Lernzuwachs – IG1 vs. KG.....	351
Tabelle 348: Kodiermanual – Literaturbezug – Test auf Normalverteilung	351
Tabelle 349: Kodiermanual – Literaturbezug – Wissenszuwachs – Nichtparametrisch	351
Tabelle 350: Kodiermanual – Literaturbezug – Wissenszuwachs – Parametrisch.....	351
Tabelle 351: Kodiermanual – Literaturbezug – Wissenszuwachs – Nichtparametrisch – Feedbackgruppe.....	352
Tabelle 352: Kodiermanual – Literaturbezug – Wissenszuwachs – Nichtparametrisch – Selbstreflexionsgruppe	352
Tabelle 353: Kodiermanual – Literaturbezug – Wissenszuwachs – Nichtparametrisch – Kontrollgruppe	352
Tabelle 354: Kodiermanual – Literaturbezug – Wissenszuwachs – Parametrisch – Feedbackgruppe.....	352
Tabelle 355: Kodiermanual – Literaturbezug – Wissenszuwachs – Parametrisch – Selbstreflexionsgruppe	352
Tabelle 356: Kodiermanual – Literaturbezug – Wissenszuwachs – Parametrisch – Kontrollgruppe	353
Tabelle 357: Literaturbezug – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. IG2	353
Tabelle 358: Literaturbezug – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG	353
Tabelle 359: Literaturbezug – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2	353
Tabelle 360: Literaturbezug – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG	353
Tabelle 361: Extremgruppenvergleich – Einfluss PCK-Test auf Reflexionskompetenztest	354
Tabelle 362: Extremgruppenvergleich – Einfluss PCK-Test auf Gesamtpunktzahl des Kodiermanuals.....	354
Tabelle 363: Extremgruppenvergleich – Einfluss CK-Test auf Reflexionskompetenztest	354
Tabelle 364: Extremgruppenvergleich – Einfluss PCK-Gesamtpunktzahl Kodiermanual	354
Tabelle 365: Einfluss des Vortestergebnisses des PCK-Tests auf die Leistungsentwicklung im Reflexionskompetenztest	355
Tabelle 366: Einfluss des Vortestergebnisses des PCK-Tests auf die Entwicklung der Gesamtpunktzahl des Kodiermanuals	355
Tabelle 367: Einfluss des Vortestergebnisses des CK-Tests auf die Leistungsentwicklung im Reflexionskompetenztest	355
Tabelle 368: Einfluss des Vortestergebnisses des CK-Tests auf die Entwicklung der Gesamtpunktzahl des Kodiermanuals	356

10.2. Entwickelte Testinstrumente

10.2.1. Reflexionskompetenztest

1. Welche Art der Reflexion verspricht die besten und vielfältigsten Nutzungsmöglichkeiten?

Mündliche Reflexion	☐
Reflexionsbögen	☐
Portfolioeinträge	☐
Tonband- oder Diktiergeräte	☐
Videografie	☒

2. Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu:

„An dieser Stelle wäre das forschend-entwickelnde Verfahren für die Schüler interessanter gewesen. Dafür spricht, dass die Motivation und das Interesse für Chemie mit jedem Versuch, den die Schüler selbst entwickeln, steigt. Manchmal wird jedoch in einzelnen Gruppen das freie Arbeiten zu sehr ausgenutzt und sie erreichen nichts, machen nur Blödsinn. Bei dieser Gruppe habe ich mich deswegen für ein sogenanntes „Kochrezept“ entschieden. Das würde ich beim nächsten Mal evtl. ändern.“

„descriptive reflection“ (beschreibende Reflexion)	☐
„descriptive writing“ (Beschreibung)	☐
„critical reflection“ (kritische Reflexion)	☐
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☒

3. Die von der Kultusministerkonferenz der Länder [KMK] formulierten Lehrerbildungsstandards (2014) unterscheiden vier Kompetenzbereiche. Welcher aufgeführte Punkt gehört nicht dazu?

Kompetenzbereich Unterrichten	☐
Kompetenzbereich Erziehen	☐
Kompetenzbereich Innovieren	☐
Kompetenzbereich Beurteilen	☐
Kompetenzbereich Evaluieren	☒

4. Die KMK formulierte Lehrerbildungsstandards (2014), die verschiedene Domänen des Professionswissens von Lehrkräften unterscheiden. Welche der aufgeführten Kompetenzen beinhaltet unter anderem die Reflexionskompetenz?

Lehrerinnen und Lehrer planen Unterricht fach- und sachgerecht und führen ihn sachlich und fachlich korrekt durch.	☒
Lehrerinnen und Lehrer erfassen die Leistungsentwicklung von Schülerinnen und Schülern und beurteilen Lernen und Leistungen auf der Grundlage transparenter Beurteilungsmaßstäbe.	☐

10. Anhang

Lehrerinnen und Lehrer fördern die Fähigkeiten von Schülerinnen und Schülern zum selbstbestimmten Lernen und Arbeiten.	☐
Lehrerinnen und Lehrer kennen die sozialen und kulturellen Lebensbedingungen von Schülerinnen und Schülern und nehmen im Rahmen der Schule Einfluss auf deren individuelle Entwicklung.	☐
Lehrerinnen und Lehrer finden Lösungsansätze für Schwierigkeiten und Konflikte in Schule und Unterricht.	☐

5. Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu:

„Mein Ziel war hier eigentlich, mich zurückzunehmen und die Schüler im Sinne des Naturwissenschaftlichen Erkenntnisweges selbst auf die richtige Theorie kommen zu lassen. Um dies zu erreichen muss mehr geschaut werden, alle Schüler – egal wie fit oder motiviert – ins Boot zu holen. Manche Schüler motiviert ein auftretendes Problem zu Höchstleistungen. Andere sind schnell frustriert und geben auf. Ziel ist es in jedem Fall, frühzeitig zu erkennen, wer Hilfe benötigt und wer evtl. dann doch selbst zu dem gewünschten Ergebnis kommt.“

„descriptive reflection“ (beschreibende Reflexion)	☐
„descriptive writing“ (Beschreibung)	☐
„critical reflection“ (kritische Reflexion)	☒
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☐

6. Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu:

„Dieser Versuch ist als Schülerexperiment im Microscale-Maßstab geeignet.“

„descriptive reflection“ (beschreibende Reflexion)	☐
„descriptive writing“ (Beschreibung)	☒
„critical reflection“ (kritische Reflexion)	☐
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☐

7. Das ALACT-Modell nach Korthagen beschreibt einen Reflexionskreislauf in folgenden fünf Phasen:

1. Handlung 2. Rückblick auf die Handlung 3. Bewusstsein der essenziellen Aspekte 4. Kreieren alternativer Handlungsformen 5. Veränderte Handlung	☒
1. Handlung 2. Rückblick auf die Handlung 3. Kreieren alternativer Handlungsformen 4. Bewusstsein der essenziellen Aspekte 5. Veränderte Handlung	☐
1. Handlung 2. Bewusstsein der essenziellen Aspekte 3. Rückblick auf die Handlung	☐

4. Kreieren alternativer Handlungsformen 5. Veränderte Handlung	
1. Handlung 2. Kreieren alternativer Handlungsformen 3. Rückblick auf die Handlung 4. Bewusstsein der essenziellen Aspekte 5. Veränderte Handlung	☐
1. Handlung 2. Kreieren alternativer Handlungsformen 3. Bewusstsein der essenziellen Aspekte 4. Rückblick auf die Handlung 5. Veränderte Handlung	☐

8. Das zentrale Thema zweier Phasen im ALACT-Modell nach Korthagen ist die Selbstbetrachtung. Ziel ist Ortung und Ordnung der eigenen Gedanken und Gefühle, bisherige Verhaltensweisen und was sie bewirkt haben. Welche zwei Phasen sind gemeint?

„Handlung“ und „Rückblick auf die Handlung“	☐
„Rückblick auf die Handlung“ und „Bewusstsein der essenziellen Aspekte“	☒
„Bewusstsein der essenziellen Aspekte“ und „Kreieren alternativer Handlungsformen“	☐
„Kreieren alternativer Handlungsformen“ und „Veränderte Handlung“	☐
„Handlung“ und „Veränderte Handlung“	☐

9. Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu:

„Es stellte sich heraus, dass dieser Schüler neu in der Klasse ist und in der alten Schule keine Versuche gemacht wurden. Ein wünschenswertes Ziel wäre, die Kompetenzstandards in jeder Schule zwingend umzusetzen. Die Vergleichbarkeit hat oberste Priorität. Das Ziel ist doch, Schüler gerecht und auf gleichem Niveau in ganz Deutschland zu vergleichen und zu benoten. Eine bestimmte Note sollte eine bestimmte Aussage über das Können eines Schülers vermitteln und nicht je Schule unterschiedlich viel wert sein.“

„descriptive reflection“ (beschreibende Reflexion)	☐
„descriptive writing“ (Beschreibung)	☐
„critical reflection“ (kritische Reflexion)	☒
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☐

10. Welches ist keines der Reflexionsniveaus nach Hatton & Smith (1995)

„descriptive reflection“ (beschreibende Reflexion)	☐
„descriptive writing“ (Beschreibung)	☐
„critical reflection“ (kritische Reflexion)	☐
„critical writing“ (kritische Beschreibung)	☒
„dialogic reflection“ (dialogische Reflexion)	☐

10. Anhang

11. Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu:

„Bei der Hypothesenbildung tat sich die Gruppe 3 sehr schwer. Hier wäre ein Tipp von mir wohl hilfreich gewesen.“

„descriptive reflection“ (beschreibende Reflexion)	☒
„descriptive writing“ (Beschreibung)	☐
„critical reflection“ (kritische Reflexion)	☐
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☐

12. Neben der Reflexionsbreite und der Reflexionstiefe existiert nach Niggli (2005) noch eine dritte Dimension:

Reflexionshöhe	☐
Reflexionsebene	☒
Reflexionsgespräch	☐
Reflexionsbereich	☐
Reflexionsabstand	☐

13. Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu:

„Am Ende des Versuchs, sollen die Schüler ihre Apparaturen, soweit möglich, selbst reinigen und aufräumen.“

„descriptive reflection“ (beschreibende Reflexion)	☐
„descriptive writing“ (Beschreibung)	☒
„critical reflection“ (kritische Reflexion)	☐
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☐

14. Welcher Inhalt gehört, ausgehend von der Beschreibung der PCK Inhalte nach Park & Oliver (2007), laut Wischmann nicht zur Reflexionsbreite?

Selbstwirksamkeit	☐
Vermittlungsstrategien	☐
Evaluation	☐
Modelleinsatz	☒
Schülervorstellung	☐

15. Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu:

„Es sollte am Ende bei jedem Schüler die Sicherung im Heft stehen und auch verstanden worden sein. In diesem Fall wäre eine Differenzierung sinnvoll gewesen. Der Niveauunterschied war deutlich zu spüren. Dieser Schüler bräuchte evtl. eine individuellere Förderung und mehr Übung darin, mit anderen zu kooperieren. Er missbrauchte die Spritzflasche als Wasserpistole. So konnte diese Gruppe mein angedachtes Lernziel nicht erreichen und der Merksatz machte dann so gar keinen Sinn.“

„descriptive reflection“ (beschreibende Reflexion)	☐
„descriptive writing“ (Beschreibung)	☐
„critical reflection“ (kritische Reflexion)	☒
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☐

16. Die „Reflexion in der Handlung“ verläuft in sechs Phasen und wird in der ersten Phase durch eine problematische Situation ausgelöst. Welche der folgenden fünf Phasen ist nach Park & Oliver (2007) nicht „in der Handlung“ möglich?

Problemdefinition erfolgt in einem ersten Schritt durch Analogieschlüsse und ‚Benennen‘.	☐
Der nächste Schritt umfasst nun die Verwirklichung der ersten Problemdefinition.	☐
Die handelnde Verwirklichung der ersten Problemdefinition stellt gleichzeitig die experimentelle Prüfung dar.	☐
Bei diesem Experiment wird die hypothetische Definition der gegebenen Situation übergestülpt.	☐
Das Wissen, welches hinter der professionellen Tätigkeit steht, kann für andere Personen transparent gemacht werden.	☒

17. Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu:

„Bei der Versuchsskizze habe ich mich für die Tafel entschieden, da ich dachte, ich könnte Schüler nach vorne holen und zeichnen lassen. Doch der Hofmann'sche Zersetzungsapparat war eine Stufe zu schwer. In dem Fall wäre ein Arbeitsblatt mit abgedruckter Skizze Zeit sparender und vor allem ordentlicher gewesen.“

„descriptive reflection“ (beschreibende Reflexion)	☐
„descriptive writing“ (Beschreibung)	☐
„critical reflection“ (kritische Reflexion)	☐
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☒

18. Baumert & Kunter (2006, S. 481f.) stellen ein Modell für die professionelle Kompetenz von Lehrkräften vor. Gemäß diesem Modell entsteht professionelle Handlungskompetenz aus dem Zusammenspiel verschiedener Komponenten. Welche Komponente gehört nicht dazu?

Überzeugungen/ Werthaltungen	☐
Fähigkeit zur Distanzierung	☒
Professionswissen	☐
Motivationale Orientierungen	☐
Selbstregulative Fähigkeiten	☐

19. Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu:

10. Anhang

„In dem Fall hätte ein Becherglas wohl auch gereicht! So war es bei manchen wohl etwas zu viel Substanz.“

„descriptive reflection“ (beschreibende Reflexion)	☒
„descriptive writing“ (Beschreibung)	☐
„critical reflection“ (kritische Reflexion)	☐
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☐

20. Reflexion kann zusätzlich nach Wischmann (2015) aus vier unterschiedlichen Perspektiven erfolgen. Welche Perspektive gehört nicht dazu?

Selbstreflexion	☐
Schülerperspektive	☐
Fachperspektive	☐
Fremdperspektive	☒
Theorieperspektive	☐

21. Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu:

„Durch den zweiten Versuch wurde die Zeit am Ende knapp. Aber für den Versuch spricht eindeutig die zusätzliche Motivation der Schüler.“

„descriptive reflection“ (beschreibende Reflexion)	☐
„descriptive writing“ (Beschreibung)	☐
„critical reflection“ (kritische Reflexion)	☐
„critical writing“ (kritische Beschreibung)	☐
„dialogic reflection“ (dialogische Reflexion)	☒

10.2.2. Kodiermanual

Kodebuch – Reflexionskompetenz

Überblick:

Das vorliegende Kodiermanual ist zur Bewertung schriftlicher Reflexionen von Chemielehramtsstudierenden über gehaltene Unterrichtsstunden konzipiert. Die Analyse beschränkt sich auf zwei große Teilbereiche (Reflexionsbreite und -tiefe) mit den jeweiligen zu kodierenden Kategorien:

1. Reflexionsbreite

- Lehrerperformanz
- Lehrerverhalten
- Vermittlung von chemischem Fachwissen
- Regeln im Chemieunterricht
- Schüleraktivierung
- Sozialform
- Schülerexperiment
- Organisationsform
- Medien

j. Sonstiges

2. *Reflexionstiefe*

- a. Beschreibung
 - i. Rechtfertigung
 - ii. Abwägen
 - iii. Zielformulierung
- b. Bewertung (positiv/negativ)
 - i. Rechtfertigung
 - ii. Abwägen
 - iii. Zielformulierung
- c. Alternative
 - i. Rechtfertigung
 - ii. Abwägen
 - iii. Zielformulierung
- d. Verbesserungsvorschlag/Konsequenz
 - i. Rechtfertigung
 - ii. Abwägen
 - iii. Zielformulierung
- e. Perspektive
 - i. Schülerperspektive
 - ii. Fachperspektive
 - iii. Theorieperspektive
 - iv. Politisch/Soziale Perspektive
- f. Literaturbezug

Aufbau des Kodebuchs:

Zunächst werden allgemeine Kodieranweisungen (KA) festgelegt.

Im Anschluss erfolgt die genaue Ausformulierung der drei Durchgänge eines Messzeitpunktes. Hierbei werden die zwei großen Teilbereiche, Reflexionsbreite und -tiefe durch Festlegung der drei Analyseeinheiten definiert. Das sind:

Kodiereinheit: Kleinster Materialbestand, welcher ausgewertet werden darf (minimaler Textteil).

Kontexteinheit: Größter Materialbestand, welcher ausgewertet werden darf (maximale Texteinheit).

Auswertungseinheit: Entspricht der Reihenfolge der auszuwertenden Textteile.

Die wichtigsten allgemeinen Kodieranweisungen werden erneut aufgeführt.

Die zu kodierenden Kategorien sind wie folgt unterteilt:

- 1. Definition der Kategorien
 - a. Allgemein
 - b. Differenzierung
- 2. Ankerbeispiele
 - a. Allgemein

- b. Differenzierung (optional)
- 3. Kodierregeln
 - a. Allgemein
 - b. Differenzierung (optional)

Die jeweiligen Differenzierungen dienen als Erklärungen oder Abgrenzungen zu zuvor aufgestellten Beispielen oder Kodierregeln.

Allgemeine Kodieranweisungen (KA):

Ein Messzeitpunkt, also ein Textdokument, wird insgesamt dreimal durchgegangen.

1. Durchgang des Materials: Kodierung der Reflexionsbreite

KA1: Zunächst wird das Material in die zehn Kategorien der Reflexionsbreite kodiert.

KA2: Es werden Absätze nach Sinnzusammenhang gebildet. Diese bestehen aus mindestens einem der folgenden Punkte:

Reinen Beschreibungen eines Teils einer Unterrichtsstunde, den zugehörigen Bewertungen, ob die Probanden zufrieden waren, der Anführung von Gründen für die Einschätzung, der Nennung von Zielen, einem Abwägen der getroffenen Entscheidungen, der Nennung von Alternativen oder einem Ziehen von Konsequenzen aus den vorherigen Überlegungen oder der Nennung von Verbesserungsvorschlägen.

Wichtig ist der stete Bezug zu einer Breitenkategorie.

KA3: Eine Kategorie wird nach dem **Aspekt, welcher vom Proband als verbesserungswürdig erachtet wird, oder nach dem Aspekt, mit welchem der Proband zufrieden war**, kodiert. Die Breite wird also nicht nach dem Ziel, der Rechtfertigung oder den Verbesserungsvorschlägen etc. ausgewählt.

KA4: Ist eine eindeutige Zuordnung nicht möglich, wird diejenige Kategorie gewählt, welche vorrangig im Absatz genannt wird.

KA5: Ist eine Zuordnung immer noch nicht möglich, wird die Kategorie Sonstiges kodiert.

2. Durchgang des Materials: Kodierung der Reflexionstiefe I

KA6: Zunächst werden nur die Punkte a) bis d) der Reflexionstiefe kodiert, einschließlich der jeweils zugehörigen Unterteilungen Rechtfertigung, Abwägen und Zielformulierungen.

KA7: Es wird nach Abschnitten der Reflexionsbreite vorgegangen.

KA8: Zur Auswertung ist der Kontext des Absatzes miteinzubeziehen.

So ist am Anfang eines Satzes nicht immer klar, ob die beschriebene Zielformulierung eines Probanden einer Beschreibung, einer Bewertung, einer Alternative oder einem Verbesserungsvorschlag zuzuordnen ist. Erst durch den Kontext des Absatzes wird dies klar. Beispiel:

„Zu Beginn der Stunde war es mein Ziel, alle Schüler durch den Alltagsbezug zu Cola im Supermarkt zu motivieren. Das bekannte Erfrischungsgetränk eignet sich dafür sehr gut, da es sprichwörtlich in aller Munde ist.“ (KÜHPAT_A5_MZP2)

Der Absatz beginnt mit einer Zielformulierung. Erst im zweiten Satz wird die positive Bewertung deutlich. Der erste Satz wird also als Zielformulierung einer Bewertung kodiert.

3. Durchgang des Materials: Kodierung der Reflexionstiefe II

Im dritten Kodierdurchgang des Messzeitpunktes erfolgt die Kodierung der Punkte e) und f) der Reflexionstiefe. Dies entspricht den Perspektiven und dem Literaturbezug.

KA9: Erfolgt keine Kodierung, entspricht dies der Lehrerperspektive.

KA10: Wenn eine andere als die Lehrerperspektive eingenommen wird, erfolgt eine zusätzliche Kodierung (von Beschreibungen, Bewertungen, Rechtfertigungen, Abwägen und Zielformulierungen) mittels der vier Perspektiven (Schülerperspektive, Theorieperspektive, Fachperspektive oder Politisch/Soziale Perspektive).

KA11: Die Perspektiven werden nur kodiert, wenn bei Beschreibungen, Bewertungen, bei Rechtfertigungen, im Falle des Abwägens oder bei Zielformulierungen Perspektivenübernahmen erfolgen.

KA12: Der Literaturbezug wird mit Kategorie f) markiert. Somit sind Dreifachkodierungen möglich.
Beispiel:

„Diese Methodenvielfalt ist nach Hilbert Meyer ein weiteres Merkmal guten Unterrichts.“ (BIESTE_A3_MZP3)

Die Probandin rechtfertigt ihren Verbesserungsvorschlag, einen Methodenwechsel zu integrieren, mithilfe eines Literaturbezugs. Dadurch wird die Theorieperspektive eingenommen. Kodierung: Rechtfertigung, Literaturbezug, Theorieperspektive.

Kodiermanual:

Für eine bessere Lesbarkeit, wurden in den Ankerbeispielen Orthographie und Interpunktion überarbeitet.

1. Durchgang des Materials: Kodierung der Reflexionsbreite

Reflexionsbreite

Kodiereinheit (minimaler Textteil): Ein Satz.

Kontexteinheit (maximale Texteinheit): Eine komplette schriftliche Reflexion eines Messzeitpunktes.

Auswertungseinheit (Reihenfolge der auszuwertenden Textteile):

Ein Messzeitpunkt wird von Anfang bis Ende durchgearbeitet und in Sinnabschnitte gemäß den neun (bzw. zehn) Kategorien unterteilt. So

entstehen Absätze, welche komplett in der Farbe der jeweiligen Kategorie kodiert werden.

Allgemein gilt:

Eine Kategorie wird nach dem zu verbessernden Punkt kodiert oder nach dem als gelungen beschriebenen Punkt und nicht nach dem Ziel oder der Rechtfertigung etc. Ist dies nicht eindeutig, wird die Kategorie gewählt, welche vorrangig im Absatz genannt ist.

Ankerbeispiele:

„Die Erklärung zur Funktionsweise der Aktivkohle verlief etwas zäh. Dem könnte man entgegenwirken, indem man ein Modell für den Adsorptionsprozess aufzeigt.“ (PAT KÜH_A5_MZP2)

Verbesserung: Erklärung verbessern

Mittel zum Zweck: Modelleinsatz

→ Die Entscheidung, welcher Breitenpunkt gewählt wird, basiert auf dem Grund, was verbessert werden soll und nicht, mit welchen Hilfsmitteln dies erreicht wird.

1. Leistung der Lehrperson/Lehrerperformanz

1. Definition

Allgemein:

Gedanken zu getroffenen Entscheidungen und der Umsetzung der Lehrkraft in den Bereichen

- Organisation vor oder im Verlauf der Stunde
- Erklärungen, Fragestellungen und Antworten der Lehrperson
- Überblick über die Klasse

Perspektive: meist in Lehrerperspektive formuliert

Differenzierung:

Lehrerleistung in der Organisation der Unterrichtseinheit:

Zeitmanagement, Versuchsvorbereitungen, praktische Vorbereitungen, Leistung im Einsatz von Visualisierungen, Austeilen/Einsammeln von Gegenständen

Lehrerleistung im Geben von Erklärungen, Fragestellungen und Antworten der Lehrperson, Art und Weise der Erklärungen oder Fragestellungen der Lehrperson, Beispiele geben, die Art und Weise des Anleitens von Übungen und der Antworten auf Fragen, stille Impulse, Art und Weise des Anleitens zum Ausfüllen von Tabellen und Lückentexten und Hefteinträgen, Art und Weise im Überleiten, Hinführen zu Themen Überblick über die Klasse, Einschätzung des Lernstandes der SuS, Einschätzungen des Vorwissens, Einschätzung der vorhandenen Basiskonzepte

2. Ankerbeispiele (AB)

Allgemein:

AB1: „Aufgrund des Zeitmangels am Ende der Stunde waren einige Schüler während der kurzen Nachbesprechung noch damit beschäftigt. Dieser Aspekt könnte mit einem besseren Zeitmanagement auch verbessert werden. So dass alle am Ende die wesentlichen Punkte auf ihrem Arbeitsblatt eintragen können und eine gemeinsame Sicherung erfolgt.“ (BIESTE A3_MZP2)

AB2: „Dies ist mir bereits bei der zum Einstieg verwendeten Fragestellung: „Sind 2-Cent Münzen vollständig aus Kupfer“ aufgefallen. Die Fragestellung hätte besser erklärt werden können bzw. die SuS mehr in die Planungsphase mit einbezogen werden können. Dadurch ist die Fragestellung nicht nur eine lebensnahe zudem sind die SuS auch aktiv an der Planungsphase einbezogen, was eine erhöhte intrinsische Motivation in der Erarbeitungsphase bewirken kann. Dadurch erkennen die SuS nämlich einen Sinn hinter ihren Experimenten und experimentieren nicht einfach stumpf vor sich hin, weil es der Lehrkörper vor schreibt, ohne den Weg bzw. das Ziel zu kennen.“ (BOXALE_A1_MZP1)

Differenzierung:

AB1: Problem war ein schlechtes Zeitmanagement. Die Folge der Verbesserung ist die zusätzliche Zeit, das Arbeitsblatt auszufüllen.

AB2: Der Proband ist nicht zufrieden mit seiner Fragestellung. Er will diese verbessern, indem er die SuS miteinbezieht.

3. Kodierregeln (KR)

Allgemein:

KR1: Wenn die Leistung der Lehrperson im Fokus von Beschreibungen, Bewertungen oder von Verbesserungsvorschlägen steht, wird der Absatz mit der ersten Breitenkategorie markiert.

KR2: Der Absatz inkludiert zugehörige Beschreibungen, Alternativen, Verbesserungsvorschläge und die jeweiligen Rechtfertigungen, Punkte des Abwägens und Zielformulierungen, auch wenn diese thematisch einem anderen Punkt zugeordnet werden müssten.

KR3: Der Absatz endet, wenn thematisch ein neuer Punkt beschrieben, bewertet oder verbessert wird.

Differenzierung:

KR4 Zeitmanagement im Schülerexperiment

→ Kategorie: Schülerexperiment

AB3: „Das Herstellen der M&M's Farben nahm sehr viel Zeit in Anspruch und es hätten mehr Behälter für die entfärbten M&M's bereitgestellt werden müssen. Um das Ziel zu erreichen, dass die Schüler die Grundzüge der Chromatographie anhand des Schülerexperiments erklären, wäre es vielleicht besser, wenn die M&M's Farben bereits vorbereitet sind, somit würde mehr Zeit für das eigentliche Experiment verbleiben. Auch sollte die Versuchsbeschreibung noch etwas deutlicher erklären, wo die Farben der M&M's auf das Filterpapier zu platzieren sind. Des Weiteren sollte schon während des Experiments darauf eingegangen werden, dass das Filterpapier die stationäre Phase und das Wasser die mobile Phase ist.“ (BERCOR_A6_MZP1)

KR5 Fachlich falsche Erklärungen der Lehrperson sind abzugrenzen von unsauberen Erklärungen → Kategorie: Vermittlung von chemischem Fachwissen

AB4: „Ein weiterer Kritikpunkt ist, dass ich das „leichter werden“ des schlechten Eis leider fachlich komplett falsch erklärt habe, an dieser Stelle muss ich mich beim nächsten Mal fachlich besser vorbereiten und zudem genauer darauf achten, wie die SuS ihre Antwort formulieren und bei Fehlvorstellungen bzw. falschen Antworten eingreifen, damit sich die Fehlvorstellungen nicht festigen bzw. einschleichen können. Der Pro-

	zess des schlechtwerdenden Eis hätte man auch gut mit einem Arbeitsblatt und Informationstext besprechen können, dadurch wäre die Stunde auch nicht bereits nach 30 Minuten zu Ende gewesen und es hätte der Stunde einen wesentlich besseren roten Faden verliehen.“ (BOXALE_A1_MZP3) KR6 Inhalt schlechter Lückentexte und Tabellen auf z. B. Arbeitsblättern ist abzugrenzen von schlechten Hinführungen/Erklärungen zu den Lückentexten/Tabellen → Kategorie: Medien AB5: „Das Problem, dass das Arbeitsblatt nicht richtig ausgefüllt wurde, könnte man durch eine konkrete Angabe möglicherweise unterbinden. Dennoch sollte das Ziel das selbstständige Arbeiten der Schüler in Gruppen so beibehalten werden, hierfür wäre eine Zeitangabe und der konkrete Auftrag, wann welche Aufgabe erledigt werden sollte, zu Beginn sinnvoll.“ (BIESTE_A3_MZP1) KR7 Guter/Schlechter Einstieg auf Inhalt bezogen und nicht auf Lehrerperformanz → Kategorie: Unterrichtsverlauf AB6: „Als Einleitung der Stunde habe ich die Problemstellung verwendet, dass es grüne Smarties gibt, jedoch auf der Verpackung kein grüner Farbstoff angegeben ist. Das ist zwar eine gute Hinführung zu dem Thema, da es einen Alltagsbezug hat und die Schüler mit den Smarties weiterarbeiten, aber das Problem ist schnell gelöst und die Schüler erkennen, dass grün ein Farbstoffgemisch von blau und gelb ist. Außerdem ist es kein sehr packender und interessanter Einstieg. Ein alternativer Einstieg wäre die Vorstellung des Textes „So eine Gemeinschaft“, in welchem ein Schüler durch eine Textnachricht gemobbt wird und man herausfinden will, wer das war. Durch geschickte Hinleitung seitens des Lehrers stellen die Schüler die Hypothese auf den Farbstoff des verwendeten Stiftes zu untersuchen und werden so auf das Thema der Farbtrennung (Chromatographie) aufmerksam und das Trennprinzip kann mithilfe von Smarties durchgeführt werden.“ (SIMMON_C6_MZP3) KR8 Wissensaufbau auf Basiskonzepten ist abzugrenzen vom Überblick über das Vorwissen einer Klasse → Kategorie: Vermittlung von chemischem Fachwissen

2. Lehrerverhalten	
	1. Definition Allgemein: Verhalten der Lehrperson betreffend Sprache, Mimik, Gestik, Verhalten, Begrüßung und Loben Differenzierung: Sprache: Dialekt oder Füllwörter, Lautstärke, Verständlichkeit, Lehrerecho Verhalten: Lehrperson hält sich selbst an Regeln; Vorbildfunktion (Kaugummi kauen, saubere Schrift etc.), Körperhaltung, Raumnutzung, Nervosität, offene Art.
	2. Ankerbeispiele Allgemein:

	AB1: „Sprachlich ist mir aufgefallen, dass ich insbesondere in Gedankenpausen bzw. auch in Sprechpausen oft das Füllwort „äh“ verwendet habe. Zudem wurde ich darauf aufmerksam gemacht, dass ich teilweise in den bayerischen Dialekt ver falle.“ (BOXALE_A1_MZP1) AB2: „Mit meinen Äußerungen bin ich allgemein sehr zufrieden, weil ich Lehrerecho vermieden habe und verschiedene Ansätze gesammelt habe.“ (KÜHPAT_A5_MZP1) <i>3. Kodierregeln</i> Allgemein: KR1: Wenn das Verhalten der Lehrperson im Fokus von Beschreibungen, Bewertungen oder von Verbesserungsvorschlägen steht, wird der Absatz mit der zweiten Breitenkategorie markiert. KR2: Der Absatz inkludiert <u>zugehörige</u> Beschreibungen, Alternativen, Verbesserungsvorschläge und die jeweiligen Rechtfertigungen, Punkte des Abwägens und Zielformulierungen, auch wenn <u>diese thematisch einem anderen Punkt zugeordnet werden müssten</u>. KR3: Der Absatz endet, wenn thematisch ein neuer Punkt beschrieben, bewertet oder verbessert wird. Differenzierung: KR4: Vorbildfunktion in Bezug auf den Chemieunterricht (Schutzbrille, Kittel) → Kategorie: Regeln im Chemieunterricht AB3: „Auf die Sicherheit muss ebenfalls besser geachtet werden, Sowohl ich selber als auch die SuS müssen Brillen tragen, den Kittel schließen und Kaugummis im Labor sind zu entfernen. Das sollte beim nächsten Mal deutlicher erwähnt und eingehalten werden.“ (BIESTE_A3_MZP1)

	3. Vermittlung von chemischem Fachwissen <i>1. Definition</i> Allgemein: Vermittlung von Fachwissen in Bezug auf das Fach Chemie. Differenzierung: Vermittlung durch die Lehrperson: Lernziele, Lehrerdemonstrationsexperiment, Lehrplan, fachliche Fehler Fachwissen der Schülerinnen und Schüler: Wiederholung von Gelerntem, Fachbegriffe, Apparaturen, Definitionen, Transferaufgaben mit dem Hintergrund, vernetztes Wissen aufzubauen <i>2. Ankerbeispiele</i> Allgemein: AB1: „Bei der Besprechung der Experimente am Ende der Stunde wurde nicht genannt, dass Eindampfen aufgrund unterschiedlicher Siedepunkte und Filtrieren unterschiedlicher Teilchengröße möglich ist. Zudem hat ein Schüler Sublimieren genannt

	(Kondensieren war gefragt). Um das in Zukunft zu vermeiden, sollte unbedingt geklärt werden, was man unter diesen Fachbegriffen versteht.“ (BIESTE_A3_MZP3) AB2: „Ein weiteres Ziel meiner Stunde, den Schülern die Dichte zu vermitteln ist mir leider nicht vollständig gelungen. Die SuS haben meiner Meinung nach das Archimedische Prinzip (Körper verdrängt das Volumen das er selbst besitzt) verstanden und auch, dass die Dichte eine Stoffeigenschaft ist, welche einen linearen Zusammenhang zwischen Massen und Volumen beschreibt. Was die Dichte aber konkret aussagt bzw. welche Auswirkungen sie im Alltag hat (Holz schwimmt, Schiffe schwimmen, Regulation von Tauchtiefe von Fischen/U-Booten) ist leider nicht klargeworden, da ich es schlichtweg versäumt habe, konkrete Beispiele aus dem Alltag mit einzubeziehen. All diese konkreten Alltagsbeispiele im Unterricht zu besprechen wird den zeitlichen Rahmen sprengen, es besteht aber durchaus die Möglichkeit, den SuS im Rahmen einer Hausaufgabe, mithilfe verschiedener „Tiny-URL’s“, auf (selbst erstelltes) Material/Website zu führen welche sich mit diesen Thematiken beschäftigen.“ (BOXALE_A1_MZP1) AB3: „Ein weiterer Kritikpunkt ist, dass ich das „leichter werden“ des schlechten Eis leider fachlich komplett falsch erklärt habe, an dieser Stelle muss ich mich beim nächsten Mal fachlich besser vorbereiten und zudem genauer darauf achten, wie die SuS ihre Antwort formulieren und bei Fehlvorstellungen bzw. falschen Antworten eingreifen, damit sich die Fehlvorstellungen nicht festigen bzw. einschleichen können. Der Prozess des schlechtwerdenden Eis hätte man auch gut mit einem Arbeitsblatt und Informationstext besprechen können, dadurch wäre die Stunde auch nicht bereits nach 30 Minuten zu Ende gewesen und es hätte der Stunde einen wesentlich besseren roten Faden verliehen.“ (BOXALE_A1_MZP3) <i>3. Kodierregeln</i> Allgemein: KR1: Wenn die Vermittlung von chemischem Fachwissen im Fokus von Beschreibungen, Bewertungen oder Verbesserungsvorschlägen steht, wird der Absatz mit der dritten Breitenkategorie markiert. KR2: Der Absatz inkludiert <u>zugehörige</u> Beschreibungen, Alternativen, Verbesserungsvorschläge und die jeweiligen Rechtfertigungen, Punkte des Abwägens und Zielformulierungen, <u>auch wenn diese thematisch einem anderen Punkt zugeordnet werden müssten</u>. KR3: Der Absatz endet, wenn thematisch ein neuer Punkt beschrieben, bewertet oder verbessert wird. KR4: Diese Kategorie nimmt hauptsächlich Bezug auf die Leistung der Schülerinnen und Schüler und ist klar abzugrenzen von der Lehrerperformance. KR5: Bezug nehmend auf die Lehrperson, betrifft diese Kategorie Überlegungen zu Lernzielen, zu einem Lehrerdemonstrationsexperiment, zum Lehrplan und fachliche Fehler. Differenzierung: KR6: Gute/schlechte/fehlende ergänzende Erklärung der Lehrperson → Kategorie: Leistung der Lehrperson

AB4: „Bei der Sammlung weiterer Stoffgemische wurde von einem Schüler richtig-erweise Wasser genannt. Womöglich wäre es sinnvoll, darauf noch ein mal genauer einzugehen und den Unterschied zwischen reinem Wasser und Leitungswasser zu ver-deutlichen, falls das nicht für alle Schüler klar ist.“ (BIESTE_A3_MZP2)

4. Regeln im Chemieunterricht

1. Definition

Allgemein:

Einführen, Einhalten und Umsetzen von Sicherheitsregeln, Aufträgen und Disziplin im Chemieunterricht

Differenzierung:

Einführen, Einhalten und Umsetzen von Sicherheitsregeln: Sicherheitsbelehrung (Schutzbrille, Kittel), Vorbildfunktion (Schutzbrille, Kittel)

Geben, Einhalten und Umsetzen von Aufträgen im disziplinierten Kontext: Arbeitsanweisungen, Arbeitsaufträge, Versuchsanleitungen, Aufräumanleitung, Ansagen

Einhalten und Umsetzen von Disziplin: Disziplinprobleme, lernförderliches Klima, gute Lernumgebung, Konsequenz

2. Ankerbeispiele

Allgemein:

AB1: „Das nächste Teilziel wäre gewesen, dass die Schüler die Grundzüge der Chromatographie anhand eines Schülerexperiments erklären. Vor dem Experiment sollte ich auch darauf achten, dass von allen Schülern die Schutzbrillen getragen werden und dass auch ich als Vorbild die Schutzbrille aufsetze, damit die Sicherheitsvorkehrungen eingehalten werden.“ (BERCOR_B6_MZP1)

AB2: „Bei dem Übergang zum selbstständigen Arbeiten im Schülerexperiment sind die Sicherheitshinweise und das Arbeitsblatt zu kurz gekommen. Ich hätte hier mehr darauf hinweisen müssen, wie der Bunsenbrenner funktioniert und welche Gefahr von ihm ausgeht, um die chaotische Situation gegen Ende der Gruppenarbeitsphase zu vermeiden. Außerdem haben viele Schüler einfach begonnen, Versuche durchzuführen, statt diese zuerst zu planen und sich Notizen auf dem Arbeitsblatt zu machen.“ (SÜSSAR_A2_MZP1)

AB3: „Die Sicherheitsunterweisung erfolgte wie üblich im Lehrervortrag und verlief ohne Probleme. Das Ziel, dass die SuS die Sicherheitsvorschriften einhielten, wurde ausnahmslos erreicht. Für den Lehrervortrag spricht hier vor allem die aus Sicht des Schülers empfundene Autorität.“ (KÜHPAT_A5_MZP3)

Differenzierung:

AB1: Kritisiert wird die fehlende Umsetzung der Sicherheitsregeln.

AB2: Kritisiert werden die fehlenden Sicherheitshinweise.

AB3: Der Bezug dieses Absatzes liegt hauptsächlich auf den Sicherheitsvorschriften und nicht auf dem Lehrervortrag.

	<i>3. Kodierregeln</i> Allgemein: KR1: Wenn Regeln des Chemieunterrichts im Fokus von Beschreibungen, Bewertungen oder Verbesserungsvorschlägen stehen, wird der Absatz mit der vierten Breitenkategorie markiert. KR2: Der Absatz inkludiert <u>zugehörige</u> Beschreibungen, Alternativen, Verbesserungsvorschläge und die jeweiligen Rechtfertigungen, Punkte des Abwägens und Zielformulierungen, <u>auch wenn diese thematisch einem anderen Punkt zugeordnet werden müssten</u>. KR3: Der Absatz endet, wenn thematisch ein neuer Punkt beschrieben, bewertet oder verbessert wird. Differenzierung: KR4: Schüler nehmen aktiv am Unterrichtsgeschehen teil, mit größtmöglicher aktiver Lernzeit. → Kategorie: Schüleraktivierung KR5: Vorbildfunktion nicht auf die Sicherheit, sondern auf das Verhalten bezogen → Kategorie: Lehrerverhalten KR6: Ungenaue Erklärungen und vage Aussagen, ohne disziplinären Kontext. → Kategorie: Lehrerperformanz

	5. Schüleraktivierung <i>1. Definition</i> Allgemein: Aktivierung der Schülerinnen und Schüler, um eine maximale aktive Lernzeit zu ermöglichen Differenzierung: Aktive Lernzeit: Differenzierung, Zusatzaufgaben, Lernprozesse, eigene/gemeinsame Erarbeitung, SuS aktivieren eigenes Vorwissen Aktive Teilnahme: Wertung von Lehrer-Schüler-Gesprächen, SuS-Antworten, Schülerverhalten, Präsentation/Besprechung von Ergebnissen, Aktivierung der SuS <i>2. Ankerbeispiele</i> Allgemein: AB1: „Bei dem letzten Teilziel, die Schüler nennen verschiedene Arten der Chromatographie, wurde nur ein Vortrag gehalten. Die Schüler sollten besser mit eingebunden werden oder es sollten Rückfragen gestellt werden, um sicher zu gehen, dass die Schüler verschiedene Chromatographie-Methoden auch nennen können.“ (BERCOR_B6_MZP1) AB2: „Einige Schüler waren mit dem Trennen der Stoffgemische schon fertig, während andere noch am Arbeiten waren. Dieses Problem könnte man lösen, indem man differenziert und den schnelleren Schülern Zusatzaufgaben gibt.“ (BIESTE_A3_MZP1) Differenzierung: AB1: Kritisiert werden die Umsetzung im Vortrag und der fehlende Einbezug der SuS, also die fehlende Aktivierung der SuS. AB2: Kritisiert werden die ungleiche Arbeitsgeschwindigkeit und die dadurch auftretende fehlende Differenzierung. <i>3. Kodierregeln</i> Allgemein: KR1: Wenn die Aktivität der Schülerinnen und Schüler im Fokus von Beschreibungen, Bewertungen oder von Verbesserungsvorschlägen steht, wird der Absatz mit der fünften Breitenkategorie markiert. KR2: Der Absatz inkludiert <u>zugehörige</u> Beschreibungen, Alternativen, Verbesserungsvorschläge und die jeweiligen Rechtfertigungen, Punkte des Abwägens und Zielformulierungen, auch wenn <u>diese thematisch einem anderen Punkt zugeordnet werden müssten</u>. KR3: Der Absatz endet, wenn thematisch ein neuer Punkt beschrieben, bewertet oder verbessert wird.

	6. Organisationsform <i>1. Definition</i>

	Allgemein: Wechselwirkung aus Methoden, Sozialformen und Unterrichtsverfahren des Lernens bzw. des Unterrichtens, Auswahl der richtigen Methode zur Erreichung eines Lernziels. Differenzierung: Sozialform: Einzelarbeit, Partnerarbeit, Gruppenarbeit, Klassenunterricht/Plenum, Frontalunterricht Unterrichtsverfahren: forschender Unterricht, nacherfindender Unterricht, forschend-entwickelnder Unterricht, Stationenlernen, instruktionsbestimmter Unterricht, Lehrervortrag Methoden: Think-Pair-Share-Methode, Gruppenpuzzle, Mindmap Auswirkungen: Sozialkompetenz, Selbstmanagement, Methodenvielfalt
	<i>2. Ankerbeispiele</i> Allgemein: AB1: „Anschließend erfolgte die Überleitung zum Einstieg in das Experimentieren mit Bildern von weiteren Stoffgemischen, also einem verschmutzten Sandstrand. Damit sollte der Alltagsbezug geschaffen werden und Interesse geweckt werden. Den Schülern wurden dann die einzelnen Bestandteile des Stoffgemisches gezeigt. Um das Arbeiten noch offener zu gestalten, wäre eine Möglichkeit, die Inhaltsstoffe nicht vorzugeben, sondern die Schüler selbst herausfinden zu lassen, welche Inhaltsstoffe dem Stoffgemisch zugesetzt sind. Durch die Bilder des verschmutzten Sandes wird diese Aufgabe möglicherweise ohnehin erleichtert und könnte so auch für alle zu schaffen sein.“ (BIESTE_A3_MZP2) AB2: „Insbesondere zu Beginn dieses Semesters habe ich den Frontalunterricht oftmals stark favorisiert. Mittlerweile bin ich durch die Ausarbeitung dieser offenen Stunde und eines weiteren offenen gestalteten Schülerversuches im Rahmen des Seminars CSV gemerkt, dass ich durchaus in der Lage bin, auch offene Stunden einigermaßen strukturiert zu planen und durchzuführen, dass ich mich in diesem Bereich allerdings noch stark verbessern muss. Als Fazit kann ich feststellen, dass ich gelernt habe, wie wichtig es ist, ein großes Methodenrepertoire als Lehrer zu besitzen und an den jeweiligen Stellen die richtige Methode auszuwählen. Insbesondere bei fachlich komplexen Stunden, kann auch ein gut strukturierter und Abwechslungsreich gestalteter Frontalunterricht bei dem die SuS dennoch aktiviert werden, das richtige Mittel der Wahl sein.“ (BOXALE_A1_MZP3) AB3: „Zuletzt möchte ich noch kurz auf die Strukturierung der Stunde eingehen. Laut Hilbert Meyer kann man diese vor allem durch einen Wechsel der Sozialformen erreichen. Das führt auch zu einer besseren Orientierung für die SuS. Alternativ hätte man auch zu Beginn der Stunde eine Grobgliederung anführen können. Gegen diese Alternative spricht allerdings, dass die Spannung auf das weitere Vorgehen verloren gehen könnte. Beim Einstieg erreichte ich durch den Lehrervortrag und den Unterrichtsgespräch Interessiertheit bei den SuS. Beim Wechsel zum reinen Lehrervortrag war den Schülern dann klar, dass ein weiterer Schritt stattfand, die Sicherheitsunterweisung. Auch die beiden Versuche grenzten sich durch ihre jeweilige Sozialform voneinander ab und untergliederten die Stunde. Somit konnte ich auch das Ziel erreichen, dass die

	SuS den roten Faden während der Stunde nie verloren, ihre Spannung und Motivation durch Abwechslung allerdings aufrechterhalten wurde.“ (KÜHPAT_A5_MZP3) AB4: „Dabei habe ich mich dazu entschieden, die Informationen im Lehrervortrag zu referieren. Eine Alternative wäre hierfür ein Arbeitsblatt mit den Informationen gewesen, welches sich die Schüler in Einzelarbeit erarbeiten hätten können. Für den Lehrervortrag spricht die Zeitersparnis. Außerdem wollte ich den SuS nicht noch einen weiteren Text anmaßen. Dagegen spricht natürlich, dass durch das Arbeitsblatt ein Methodenwechsel stattgefunden hätte und die Schüler ihre Kompetenz geschult hätten, Informationen aus Sachtexten zu erschließen.“ (KÜHPAT_A5_MZP3) Differenzierung: AB1: Zunächst erfolgt eine Beschreibung des Einstiegs. Wünschenswertes Ziel ist dann jedoch, das Arbeiten noch offener zu gestalten. Somit wird dieser Absatz der sechsten Kategorie zugeordnet. Die Beschreibung am Anfang führt lediglich zum Thema. AB3: Der Proband schreibt zwar im ersten Satz von Strukturierung der Stunde, der Absatz handelt jedoch von der Methodenvielfalt. Hier wird also nicht auf Kategorie acht entschieden, welche die Strukturierung der Stunde mit Einstieg, Erarbeitung, Sicherung etc. beinhaltet, sondern auf Kategorie sechs. AB4: Inhalt des Absatzes ist der Vergleich Lehrervortrag mit SuS erarbeiten Informationen selbst. Hilfsmittel für Letzteres ist ein Arbeitsblatt. <i>3. Kodierregeln</i> Allgemein: KR1: Wenn Methoden, Unterrichtsverfahren bzw. Organisationsformen im Fokus von Beschreibungen, Bewertungen oder Verbesserungsvorschlägen stehen, wird der Absatz mit der sechsten Breitenkategorie markiert. KR2: Der Absatz inkludiert <u>zugehörige</u> Beschreibungen, Alternativen, Verbesserungsvorschläge und die jeweiligen Rechtfertigungen, Punkte des Abwägens und Zielformulierungen, <u>auch wenn diese thematisch einem anderen Punkt zugeordnet werden müssten</u>. KR3: Der Absatz endet, wenn thematisch ein neuer Punkt beschrieben, bewertet oder verbessert wird.

	7. Schülerexperiment <i>1. Definition</i> Allgemein: Betrifft alle Bestandteile eines Experiments, ab Start der Durchführung bis zum Aufräumen mit den Vor- und Nachteilen eines Schülerexperiments und der Strukturierung der Experimentierphase. Differenzierung: Schülerexperiment: Zeitangabe, Durchführung, Arbeitsphase, Aufbau Apparatur, Aufräumen, Materialienliste

	Auswirkungen: selbstständiges/freies/offenes Experimentieren, Materialienverbrauch, Experimentierkompetenz, naturwissenschaftlicher Erkenntnisgewinnungsprozess 2. Ankerbeispiele Allgemein: AB1: „Mein Ziel war, dass die SuS die Möglichkeit haben frei zu experimentieren und eigene Ideen umzusetzen. Daher habe ich auch mehrere Optionen offengelassen (Zusätzliche Materialien, Dekantieren oder Sieben des Sandes bzw. Abdampfen mit Reagenzglas oder Löffel) und keine Reihenfolge zur Trennung vorgegeben. Um das freie Experimentieren nicht zu stören, habe ich bewusst nicht in den Arbeitsprozess eingegriffen, auch wenn dieser nicht immer von zielführenden Handlungen geprägt war.“ (BIESTE_A3_MZP2) AB2: „Die Zeitangabe von 15 min., die ich vor der Gruppenarbeitsphase angegeben habe, wurde von mir nicht eingehalten, was dazu geführt hat, dass ich die Stunde nicht wie geplant zu Ende geschafft wurde und die Sicherung darunter gelitten hat. Eine Möglichkeit, dies nächstes Mal besser zu machen, wäre wahrscheinlich den Schülern während der Erarbeitung mitzuteilen, wie viel Zeit noch verbleibt und nach der angegebenen Zeit konsequent abzurechnen. Dadurch könnte die Aufgabe 3 des Arbeitsblattes noch erledigt werden, die wesentlich ist, um zu verdeutlichen, dass Stoffgemische aufgrund unterschiedlicher Eigenschaften voneinander getrennt werden können. So ist vielen Schülern wahrscheinlich nicht bewusst, welches Ziel mit der Anwendung verschiedener Methoden erreicht werden sollte.“ (BIESTE_A3_MZP2) AB3: „Das Herstellen der M&M's Farben nahm sehr viel Zeit in Anspruch und es hätten mehr Behälter für die entfärbten M&M's bereitgestellt werden müssen. Um das Ziel zu erreichen, dass die Schüler die Grundzüge der Chromatographie anhand des Schülerexperiments erklären, wäre es vielleicht besser, wenn die M&M's Farben bereits vorbereitet sind, somit würde mehr Zeit für das eigentliche Experiment verbleiben. Auch sollte die Versuchsbeschreibung noch etwas deutlicher erklären, wo die Farben der M&M's auf das Filterpapier zu platzieren sind. Des Weiteren sollte schon während des Experiments darauf eingegangen werden, dass das Filterpapier die stationäre Phase und das Wasser die mobile Phase ist.“ (BERCOR_B6_MZP1) Differenzierung: AB3: Als verbesserungswürdig fällt hier das Zeitmanagement während des Versuches auf. Um dies zu optimieren, wird als Mittel zum Zweck eine bessere Vorbereitung vorgeschlagen. 3. Kodierregeln Allgemein: KR1: Wenn das Schülerexperiment im Fokus von Beschreibungen, Bewertungen oder von Verbesserungsvorschlägen steht, wird der Absatz mit der siebten Breitenkategorie markiert. KR2: Der Absatz inkludiert <u>zugehörige</u> Beschreibungen, Alternativen, Verbesserungsvorschläge und die jeweiligen Rechtfertigungen, Punkte des Abwägens und Zielformulierungen, auch wenn <u>diese thematisch einem anderen Punkt zugeordnet werden müssten</u>.

	KR3: Der Absatz endet, wenn thematisch ein neuer Punkt beschrieben, bewertet oder verbessert wird. Differenzierung: KR4: Allgemeines Zeitmanagement der Stunde → Kategorie: Lehrerleistung KR5: Reine Vorbereitungen zum Experimentieren vor der Stunde → Kategorie: Lehrerleistung

8. Unterrichtsverlauf	
	<i>1. Definition</i> Allgemein: Verlauf und Strukturierung der Unterrichtsstunde Stundenzusammenfassungen/Resümee am Ende der Reflexion Differenzierung: Verlauf: Gedanken zur Planung der Stunde, Einstieg, Erarbeitungsphase, Arbeitsphasen, Zwischensicherungen, Übung, Sicherung, Puffer, Hausaufgabe Strukturierung: Roter Faden (Überleitungen, Übergänge, Hinführungen), Dynamik Stundenzusammenfassungen, Fazit
	<i>2. Ankerbeispiele</i> Allgemein: AB1: „Die Stunde wurde mit Fotos von Reinstoffen und Stoffgemischen begonnen, um die Schüler auf das Thema hinzuführen und eine Vorstellung zu geben, was unter diesen Begriffen, die im Laufe der Stunde noch weiter verwendet werden verstanden wird.“ (BIESTE_A3_MZP2) AB2: „Der Einstieg war in Ordnung, die Schüler/innen haben erklärt, dass einige Farben, wie z.B. Grün und Braun zusammengesetzte Farben sind. Der Begriff Chromatographie wurde von mir eingeführt.“ (BERCOR_B6_MZP2) AB3: „Die Sicherung der Stunde, bei der die Schüler/innen einen Lückentext zum Thema Chromatographie ausfüllen, hat ziemlich gut funktioniert. Die Schüler wussten zum großen Teil, die Lücken richtig auszufüllen. Die zwei falsch ausgefüllten Lücken, wurden von einer anderen Schülerin berichtigt. Somit hoffe ich, dass alle Schüler/innen die wesentlichen Punkte der Chromatographie verstanden haben. Wenn noch Zeit bleibt, wäre es gut, wenn die Schüler/innen die Experimentierutensilien mit aufräumen.“ (BERCOR_B6_MZP2) AB4: „Zum Schluss ist zu sagen, dass ich meine Ziele zum größten Teil realisieren konnte, aber dennoch eine Verbesserung in verschiedenen Bereichen der Stunde möglich ist.“ (CLAMAR_B5_MZP1) Differenzierung:

	AB3: Die Probandin ist zufrieden mit der Sicherung der Stunde. Als Mittel zur Sicherung wurde ein Arbeitsblatt eingesetzt. Zum Schluss des Absatzes würde sie die Sicherungsphase noch um das Aufräumen erweitern. AB4: Allgemeines Resümee der Stunde. 3. Kodierregeln Allgemein: KR1: Wenn der Unterrichtsverlauf im Fokus von Beschreibungen, Bewertungen oder von Verbesserungsvorschlägen steht, wird der Absatz mit der achten Breitenkategorie markiert. KR2: Der Absatz inkludiert <u>zugehörige</u> Beschreibungen, Alternativen, Verbesserungsvorschläge und die jeweiligen Rechtfertigungen, Punkte des Abwägens und Zielformulierungen, auch wenn <u>diese thematisch einem anderen Punkt zugeordnet werden müssten</u>. KR3: Der Absatz endet, wenn thematisch ein neuer Punkt beschrieben, bewertet oder verbessert wird. Differenzierung: KR4: Schlechte Durchführung, also Lehrerperformanz, von Überleitungen, Übergängen, Einstiegen etc. → Kategorie: Lehrerleistung

	9. Medien 1. Definition Allgemein: Kommunikationsmittel (Medium), das im Unterricht zu didaktischen Zwecken eingesetzt wird. Differenzierung: Bewertung des Medieneinsatzes wie Tafelanschrieb, Arbeitsblatt, Beamereinsatz, PPT, Taschenrechner, Handy, Film, Fotos, Modelleinsatz, Folien, Dokumentenkamera, usw. 2. Ankerbeispiele Allgemein: AB1: „Der Film zum Schluss rundete die Stunde gut ab und gab noch einmal einen Ausblick auf andere Methoden zur Entfernung des Farbstoffes. Diese ist ja Voraussetzung für die Fehlingprobe und war ein wichtiges Lernziel der Stunde.“ (KÜHPAT_A5_MZP2) AB2: „Die Ergebnissicherung war durch die Nummerierung der einzelnen Schritte übersichtlich, sodass mein Ziel, einen für die Schüler verständlichen Hefteintrag zu produzieren, mit dem sie die Stunde nacharbeiten können, erfüllt wurde.“ (SÜSSAR_A2_MZP3) Differenzierung:

	AB2: Das Ziel, einen hilfreichen Hefteintrag zu produzieren, wurde durch die Nummerierung der einzelnen Schritte erreicht. 3. <i>Kodierregeln</i> KR1: Wenn die Verwendung und Bewertung von Medien im Fokus von Beschreibungen, Bewertungen oder Verbesserungsvorschlägen steht, wird der Absatz mit der neunten Breitenkategorie markiert. KR2: Der Absatz inkludiert <u>zugehörige</u> Beschreibungen, Alternativen, Verbesserungsvorschläge und die jeweiligen Rechtfertigungen, Punkte des Abwägens und Zielformulierungen, auch wenn <u>diese thematisch einem anderen Punkt zugeordnet werden müssten</u>. KR3: Der Absatz endet, wenn thematisch ein neuer Punkt beschrieben, bewertet oder verbessert wird.

	10. Sonstiges 1. <i>Definition:</i> Allgemein: Kategorie, wenn keine Zuteilung zu anderen Breitenkategorien möglich ist Differenzierung: Gliederungen, Aufführung von theoretischen Punkten zu den verschiedensten Kategorien 2. <i>Ankerbeispiel:</i> Allgemein: AB: „Um noch einmal auf die zehn Merkmale nach Hilbert Meyer sprechen zu kommen.“ 1. Der Unterricht war im Groben recht klar strukturiert. 2. Durch die Bereitstellung der Farben war ein hoher Anteil an echter Lernzeit. 3. Ich habe mich um ein lernförderliches Klima bemüht und hatte auch den Eindruck, dass die Schüler dies merkten. 4. Inhaltliche Klarheit war meist gegeben. 5. Sinnstiftendes Kommunizieren hätte noch mehr gefördert werden können, zum Beispiel durch arbeitsteilige Gruppenarbeit zur Erarbeitung der Chromatografie-Methoden. 6. Methodenvielfalt war durch die Abwechslung von Lehrer-Schüler-Gespräch, von PowerPoint-Präsentation, Lückentext und Experiment größtenteils gewährleistet. 7. Individuelles Fördern konnte ich nicht wirklich gewährleisten, da ich die Schüler nur eine Stunde hatte. Jedoch habe ich mich bemüht, alle Schüler/innen einzubeziehen. 8. Intelligentes Üben wurde eventuell durch den Lückentext gewährleistet 9. Klare Leistungserwartung war auch meist gegeben. Die Schüler wussten, was zu tun ist und setzten die Arbeitsaufträge auch um. 10. Vorbereitete Umgebung: Ich habe mich bemüht, alle Materialien vorzubereiten.“ (BERCOR_B6_MZP3) Differenzierung:

	AB: Die Probandin führt nur ein direktes Zitat am Ende des Messzeitpunktes an, um verschiedene Breitenpunkte in der Reflexion zu belegen. Dies ist lediglich eine Ergänzung. 3. <i>Kodierregeln:</i> KR: Wenn es nicht möglich ist, den Absatz in eine der neun Kategorien der Reflexionsbreite einzuordnen, wird Kategorie 10 Sonstiges kodiert.

2. Durchgang des Materials: Kodierung der Reflexionstiefe

Reflexionstiefe <i>Kodiereinheit</i> (minimaler Textteil): Ein Teilsatz. <i>Kontexteinheit</i> (maximale Texteinheit): Ein Stichpunkt bzw. ein ganzer Satz bzw. ein Absatz der Reflexionstiefe. <i>Auswertungseinheit</i> (Reihenfolge der auszuwertenden Textteile): Pro gebildetem Absatz aus dem ersten Durchgang wird satzweise gearbeitet. Jede Teilaussage eines Satzes wird mit einer Kategorie kodiert. Eine Teilaussage kann im dritten Durchgang durch zusätzliche Perspektivenübernahme und Literaturbezüge dreifach in der Tiefe kodiert werden. Es wird im Kontext eines Abschnitts der Reflexionsbreite kodiert. Allgemein gilt: Der Kontext eines Absatzes der Reflexionsbreite wird immer miteinbezogen. Die Reflexionsbreite wird in zwei Durchgängen kodiert. 1. Durchgang: a) bis d) 2. Durchgang: e) und f)	
a)	Beschreibung
Beschreibung	1. <i>Definition</i> Allgemein: Deskriptives Schreiben Schlichte Beschreibung einer Situation ohne Wertung Allgemeine didaktische Aussagen ohne Bezug zur Stunde Beschreibung, in welcher nicht klar wird, ob sie positiv oder negativ gewertet wird. Differenzierung: Reine Wiedergabe/Beschreibung von Erlebnissen/Beobachtungen. Benennung einer Handlung. Der Autor schreibt in Erzählform ohne Wertungen, weder positive noch negative. Es werden weder Rechtfertigungen angegeben oder die Situation auf ihre Güte abgewogen, noch Ziele genannt. Beschreibungen von Verbesserungsvorschlägen oder Alternativen. Auch Beschreibungen zu Verbesserungsvorschlägen werden mit dieser Kategorie kodiert, wenn diese durch zum Beispiel Rechtfertigungen unterbrochen wurden. (Siehe Bsp. 1 oder AB5)

Sich wiederholende Punkte, welche bereits mit einer anderen Kategorie gewertet wurden, werden nicht noch einmal mit dieser kodiert, sondern als Beschreibung markiert. (Siehe AB6)

Beispiel1:

Proband macht einen Verbesserungsvorschlag.

Verbesserungsvorschlag wird gerechtfertigt.

Verbesserungsvorschlag wird genauer ausgeführt. (Siehe AB5)

Beispiel2:

Proband beschreibt, wie die Stunde unterrichtet wurde, ohne Wertung.

Proband beschreibt sich selbst in der Stunde.

Proband beschreibt, welche Aspekte oder Inhalte in der Stunde übermittelt wurden.

2. Ankerbeispiele (AB)

Allgemein:

AB1: „Nach Hilbert Meyer kann die Änderung der Sozialform verwendet werden, um die unterschiedlichen Unterrichtsphasen deutlich voneinander abzutrennen.“ (NIKLIO_A6_MZP3)

AB2: „In dieser Stunde sollte den Schülern das Verfahren der Chromatographie nähergebracht werden, indem sie einen Versuch der Trennung von Farbstoffen bei Smarties durchführen. Bei der anfänglichen Problemstellung, wie die grüne Farbe bei Smarties zustande kommt, erkannten die Schüler sofort, dass es sich um eine Mischung von blauen und gelben Farbstoffen handelt. Auch die Frage, wie man den Farbstoff von einem Smartie abtrennt, war den Schülern klar. Das Lösen als Demonstration war somit überflüssig und könnte auch weggelassen werden um Zeit zu sparen.“ (SIMMON_C6_MZP1)

AB3: „Da die Schüler für den Versuch unterschiedlich lange brauchten, sollten das nächste Mal Extrafragen auf dem Arbeitsblatt bereitstehen, mit denen sich die schnelleren Schüler beschäftigen können. Die Fragen können im Anschluss noch kurz und knapp im Plenum beantwortet werden oder als Hausaufgabe aufgegeben werden, wodurch die Schüler auch angetrieben werden, nicht zu trödeln. Eine Frage auf dem Arbeitsblatt könnte sein, andere Farbstoffgemische im Alltag zu finden.“ (SIMMON_C6_MZP2)

AB4: „Die richtige Antwort eines Schülers, dass es sich um Salz handelt, kam sehr schnell. Hier hätte ich vielleicht nicht so schnell weitergehen sollen, sondern nach Meinungen der anderen Schüler fragen, um den Einstieg nicht so schnell abzuhandeln und alle Schüler gedanklich in das Thema einzuführen.“ (SÜSSAR_A2_MZP2)

AB5: „Es hätte sich bei dem Thema angeboten, einen offeneren Unterricht zu gestalten. Dies hätte mehrere Kompetenzen der Schüler gefördert, wie die Erkenntnisgewinnung oder Kommunikation der Klasse. Ich hätte es dadurch umsetzen können, indem ich den Schülern Trennverfahren aufgezeigt hätte, auch solche die sie nicht anwenden müssten. Danach hätte ich die Stoffgemische verteilt und ihnen die Aufgabe gegeben diese zu trennen. Vorne beim Pult würden mehrere Geräte stehen, die sie zum Trennen brauchen könnten und sie sollten sich selber überlegen welche Geräte sie brauchen und diese hätten sie sich geholt. Zum Schluss hätten wir die angewendeten Methoden gesammelt und verglichen.“ (KHASEV_B3_MZP2)

	AB6: „Insgesamt wird auch das Thema der Stunde nicht wirklich klar. Spätestens nach 10 Minuten sollte der Schüler das Stundethema kennen. Bis zum Ende war nicht wirklich ersichtlich, was das Thema und das Ziel der Stunde waren.“ (KIESAN_C5_MZP1) Differenzierung: AB4: Es wird auch im Kontext nicht deutlich, ob es gut oder schlecht ist, dass der Schüler sehr schnell auf die Antwort kam. Im Anschluss wird die Lehrerperformanz überdacht. AB5: Verbesserungsvorschlag mit anschließender Rechtfertigung. Erst dann wird der Verbesserungsvorschlag genau beschrieben. AB6: Die erste negative Bewertung wird im letzten Satz lediglich wiederholt. 3. <i>Kodierregeln</i> Allgemein: KR1: Einleitungen ohne Wertung, allgemeingültige Aussagen und deskriptive Aussagen werden mit der Kategorie Beschreibung kodiert. KR2: Beschreibungen können satzweise kodiert oder mehrere aufeinander folgende beschreibende Sätze in einer Kodierung zusammengefasst werden. KR3: Wenn nicht deutlich wird, ob die Aussage positiv oder negativ ist, wird Beschreibung kodiert (vgl. AB4). Differenzierung: KR4: Um einschätzen zu können, ob die Aussage wirklich beschreibend ist, muss der Kontext miteinbezogen werden. AB4: „Einige Schüler waren mit dem Trennen der Stoffgemische schon fertig, während andere noch am Arbeiten waren. Dieses Problem könnte man lösen, indem man differenziert und den schnelleren Schülern Zusatzaufgaben gibt.“ → Der erste Satz allein ist rein deskriptiv. Erst durch den zweiten Satz wird die negative Bewertung deutlich. (Der letzte Teilsatz ist ein Verbesserungsvorschlag.)
	Rechtfertigung der Beschreibung 1. <i>Definition</i> Allgemein: Begründungen, Rechtfertigungen zu einer Beschreibung Differenzierung: Rechtfertigung der Beschreibung durch Hypothesen, Vermutung, Aufzählung einzelner Vorteile, Aufzählung einzelner Nachteile, persönliche Meinungen 2. <i>Ankerbeispiele</i> Allgemein: AB1: „Um das freie Experimentieren nicht zu stören, habe ich bewusst nicht in den Arbeitsprozess eingegriffen, auch wenn dieser nicht immer von zielführenden Handlungen geprägt war.“ (BIESTE_A3_MZP2)

	AB2: „Anschließend erfolgte die Überleitung zum Einstieg in das Experimentieren mit Bildern von weiteren Stoffgemischen, also einem verschmutzten Sandstrand. Damit sollte der Alltagsbezug geschaffen werden und Interesse geweckt werden.“ (BIESTE_A3_MZP2) Differenzierung: AB1: Der nicht erfolgte Eingriff in den Arbeitsprozess wird gerechtfertigt durch das ungestörte freie Experimentieren. AB2: Hier wird die zuvor erfolgte Beschreibung des Ablaufes begründet. <i>3. Kodierregeln</i> KR1: Mehrere folgende Rechtfertigungen werden einzeln kodiert, solange sie sich inhaltlich unterscheiden. Ansonsten wird lediglich einmal kodiert. KR2: Sich bedingende Punkte (siehe AB2 Alltagsbezug und Interesse wecken) werden lediglich einmal als Rechtfertigung kodiert.
	Abwägen der Beschreibung <i>1. Definition</i> Allgemein: Betrachtung der anderen Seite einer gerechtfertigten Beschreibung. Abwägen einer Beschreibung, bei der nicht klar wird, ob sie als positiv oder negativ gewertet wird. Differenzierung: Nach einer Rechtfertigung zu einer Beschreibung werden Vor- bzw. Nachteile gegen die zuvor getroffene Rechtfertigung genannt. Nach einer Beschreibung folgen positive und negative Argumente. <i>2. Ankerbeispiele</i> Allgemein: AB1: „Ich habe mich außerdem dazu entschlossen, das Falten des Filterpapiers zu zeigen. In der letzten Stunde hatten nämlich einige SuS damit Probleme. Allerdings nahm ich ihnen damit einen Teil des naturwissenschaftlichen Erkenntnisgewinns.“ (KÜHPAT_A5_MZP2) AB2: „Gezielt wurde dieser vor den Augen der Schüler in ihre Versuchsgefäße geschüttet, sodass nochmals deutlich wurde, dass es sich wirklich um Wein aus der Weinflasche handelt. Der Alltagsbezug wäre dadurch für die Schüler noch klarer geworden. Darauf kann man jedoch verzichten, wenn man sich die Zeitersparnis in Betracht zieht.“ (SCHKAT_A4_MZP1) AB3: „Nach Schreiber, Theyßen und Schecker lässt sich ein Experiment in Planung, Durchführung und Analyse unterteilen. Die Planung wurde den einzelnen Gruppen überlassen. Die Schüler machten sich nur wenig Gedanken über das, was sie vorhatten und probierten ihre Schritte gleich aus. Dadurch kam es vereinzelt zu Fehlschlägen. Dies ist zwar einerseits gut, da sie nun auf die Wichtigkeit einer guten Planung hingewiesen werden, andererseits kommen sie so in Zeitverzug.“ (REICHR_B2_MZP3)

	AB4: „Insgesamt bin ich was das Arbeitsblatt angeht eher zwiegespalten. Gut zum Beispiel fand ich, dass auf dem Arbeitsblatt sich eine Tabelle zum Ausfüllen und Festhalten der Ergebnisse befand, so dass die Schüler sich später an das Experiment und dessen Ergebnis nochmal erinnern können. Schlechter hingegen fand ich den Lückentext realisiert, da ich mehr das Gefühl hatte, dass dieser Text die Aufgabe einer fehlenden Erarbeitungsphase aufgriff statt einer Sicherung.“ (BADCHR_C6_MZP1) Differenzierung: AB1: Zunächst wird im ersten Satz deskriptiv berichtet. Im zweiten Satz wird dieses Verhalten gerechtfertigt. Im dritten Satz wird diese Rechtfertigung abgewogen. AB2: Zunächst wird beschrieben, dass der Wein vor den Augen in die Versuchsgläser geschüttet wurde, worauf eine Rechtfertigung folgt. Zusätzlich wird der Vorteil Alltagsbezug (Kodierung: Rechtfertigung) aufgeführt. Im Anschluss wird jedoch ein Argument gegen diese Durchführung aufgeführt, also abgewogen. AB3: Es wird aus dem Kontext nicht deutlich, ob dies positiv oder negativ bewertet wird, so werden sowohl negative als auch positive Argumente mit Abwägen kodiert. AB4: Proband ist zwiegespalten und wägt dann positiv wie negativ ab. Es wird sowohl das positive Argument als auch das negative mit Abwägen kodiert. <i>3. Kodierregeln</i> Allgemein: KR1: Pro Argument gegen zuvor beschriebene Rechtfertigungen wird mit Abwägen kodiert. KR2: Wird nach dem Abwägen wieder die erste Rechtfertigung unterstützt, wird mit Rechtfertigung zur Beschreibung kodiert. Beispiel: Beschreibung → Kodierung: Beschreibung Rechtfertigung → Kodierung: Rechtfertigung der Beschreibung Nennung Nachteil → Kodierung: Abwägen Argument für Verhalten → Kodierung: Rechtfertigung der Beschreibung KR3: Wenn nach einer Beschreibung positive und negative Argumente folgen, werden alle Argumente mit Abwägen kodiert. Eine klare Stellungnahme des Probanden erfolgt quasi nicht. Differenzierung: KR4: Wenn aus dem Kontext keine eindeutige Bewertung hervorgeht und im Anschluss sowohl positive als auch negative Argumente aufgeführt werden, werden alle Argumente, positiv wie negativ, einzeln mit Abwägen kodiert. → AB3 Zielformulierung der Beschreibung <i>1. Definition</i>

Allgemein:

Reine Zielformulierung

Differenzierung:

Formulierung eines Ziels, welches für die Unterrichtsstunde geplant war, zu erreichen und auch im Kontext rein deskriptiv ausformuliert ist.

2. Ankerbeispiele

Allgemein:

AB1: „Mein Ziel war, dass die SuS die Möglichkeit haben frei zu experimentieren und eigene Ideen umzusetzen.“ (BIESTE_A3_MZP2)

AB2: „Das Stundenziel war es, dass die SuS einige Alltagsbeispiele der Dichte erklären können und zudem die Grundlagen der Dichtebestimmung anwenden können. Insbesondere der Einstieg mit der bekannten Scherzfrage: Was ist schwerer 1kg Federn oder 1kg Blei und die nachfolgenden Beispiele (1 L Öl oder 1 L Wasser, 1 L Luft oder 1 L Wasser) hat die SuS in ihrem Alltag abgeholt.“ (BOXALE_A1_MZP2)

Differenzierung:

AB2: Im Kontext wird keine Bewertung bzw. ein rein deskriptiver Bezug zu dem formulierten Stundenziel deutlich.

3. Kodierregeln

Allgemein:

KR1: Die Formulierungen müssen eine klare Zielformulierung inkludieren.

Beispiel:

„Mein Ziel war es...“; „Ich wollte erreichen, dass...“

KR2: Es wird nur eine Zielformulierung der Beschreibung kodiert, wenn im Kontext keine Bewertung zu erkennen ist.

Differenzierung:

KR3: Wenn im Kontext eine Bewertung zu erkennen ist, wird die Kategorie Zielformulierung der Bewertung kodiert.

AB3: „Bei dem nächsten Teilziel, dass die Schüler ein Papierchromatogramm auswerten, hat es in dieser Stunde schon funktioniert, es sollte aber ein Papierchromatogramm allen Schülern gezeigt werden, bei dem zu sehen ist, wie sich die grüne Farbe in einen gelben und einen blauen Streifen auftrennen. Somit wird es für alle Schüler deutlich.“ (BERCOR_B6_MZP1)

→ Kodierung: Zielformulierung der Bewertung

KR4: Wenn anschließend ein Verbesserungsvorschlag mit Bezug zur Zielformulierung gemacht wird, erfolgt die Kodierung mit Zielformulierung des Verbesserungsvorschlags.

AB4: „Bei meinem vierten Stundenziel, der Nennung von verschiedenen Chromatographie-Methoden, sollten die Schüler/innen besser eingebunden werden.“ (BERCOR_A6_MZP2)

→ Kodierung: Zielformulierung des Verbesserungsvorschlags

	KR5: Wenn es keine klare Zielformulierung ist, wird die dem Kontext entsprechende Rechtfertigung kodiert. AB5: „Anschließend erfolgte die Überleitung zum Einstieg in das Experimentieren mit Bildern von weiteren Stoffgemischen, also einem verschmutzten Sandstrand. Damit sollte der Alltagsbezug geschaffen werden und Interesse geweckt werden.“ (BIESTE_A3_MZP2) → Kodierung: Rechtfertigung der Beschreibung
b)	Bewertung (positiv/negativ)
Bewertungen (positiv/negativ)	1. <i>Definition</i> Allgemein: Persönliche Bewertung (positiv/negativ) einer Situation Differenzierung: Bewertung einer Handlung oder einer Situation auf der Basis impliziter Annahmen und subjektiver Handlungsnormen Beispiele: Proband bewertet die Stunde/die Situation. Proband bewertet sich selbst/sein Verhalten in der Stunde. Proband bewertet Aspekte oder Inhalte, welche in der Stunde übermittelt wurden. „War gut...“; „War schlecht...“ 2. <i>Ankerbeispiele</i> Allgemein: AB1: „Einige Schüler waren mit dem Trennen der Stoffgemische schon fertig, während andere noch am Arbeiten waren. Dieses Problem könnte man lösen, indem man differenziert und den schnelleren Schülern Zusatzaufgaben gibt.“ (BIESTE_A3_MZP1) AB2: „Die Einstiegsfragen wurden von den Schülern gut beantwortet und haben sehr schnell zum Thema geführt.“ (BERCOR_B6_MZP1) AB3: „Der Einstieg führte unweigerlich zu der Problemstellung.“ (PATKÜH_A5_MZP1) AB4: „Das Problem wurde am Ende der Stunde natürlich aufgegriffen und gelöst.“ (PATKÜH_A5_MZP1) Differenzierung: AB1: Durch den Kontext im zweiten Satz wird die negative Bewertung des ersten Satzes deutlich. AB2: Hier werden die Einstiegsfragen positiv bewertet, also einfach kodiert. Die Aufzählung von zwei positiven Argumenten (gut beantwortet und schnell zum Thema geführt) ist hierbei unerheblich. AB3+4: Aus dem Kontext wird die positive Bewertung deutlich. Signalwörter in diesen Beispielen sind „unweigerlich“ bzw. „natürlich“. 3. <i>Kodierregeln</i>

	KR1: Wenn eindeutig oder aus dem Kontext eine Bewertung (positiv/negativ) einer Situation oder Handlung hervorgeht, wird die entsprechende Bewertung (positiv/negativ) kodiert. KR2: Werden zwei unterschiedliche (negative/positive) Bewertungen in einem Satz gemacht, werden beide getrennt kodiert. KR3: Es ist zu prüfen, was positiv bewertet wird, und dies wird einfach kodiert. Diese Bewertung ist unabhängig von evtl. mehrfach aufgeführten Argumenten (vgl. AB2) zur Bewertung.
	Rechtfertigung der Bewertung <i>1. Definition</i> Allgemein: Begründungen, Rechtfertigungen einer Bewertung (positiv/negativ) Differenzierung: Rechtfertigung einer positiven oder negativen Bewertung durch Hypothesen, Vermutung, Aufzählung einzelner Vorteile, Aufzählung einzelner Nachteile oder Begründung durch eine persönliche Meinung. <i>2. Ankerbeispiele</i> AB1: „Beim Einstieg waren die SchülerInnen noch relativ zurückhaltend. Möglicherweise lag das daran, dass vielen nicht klar war, was unter Reinstoffen bzw. Stoffgemischen verstanden wird.“ (BIESTE_A3_MZP1) AB2: „Aufgrund des Zeitmangels am Ende der Stunde waren einige Schüler während der kurzen Nachbesprechung noch damit beschäftigt.“ (BIESTE_A3_MZP2) <i>3. Kodierregeln</i> KR1: Mehrere folgende Rechtfertigungen einer Bewertung werden einzeln kodiert, so lange sie sich inhaltlich unterscheiden. Ansonsten wird lediglich einmal kodiert. KR2: Sich bedingende Punkte (siehe AB2 Alltagsbezug und Interesse wecken) werden lediglich einmal als Rechtfertigung kodiert.
	Abwägen der Bewertung <i>1. Definition</i> Allgemein: Betrachtung der anderen Seite einer gerechtfertigten Bewertung. Differenzierung: Nach einer Rechtfertigung zu einer getroffenen (positiven/negativen) Bewertung werden Vor- bzw. Nachteile gegen die zuvor getroffene Rechtfertigung genannt. <i>2. Ankerbeispiele</i> Allgemein:

	AB1: „Dies hat ziemlich gut funktioniert, da alle Schüler mir deren Aufmerksamkeit schenken. <i>Gegen so einen stark alltagsbezogenen Einstieg spricht allerdings, dass man sich zu sehr vom eigentlichen Thema entfernen könnte. Auch habe ich mich auf eine bestimmte Cola-Marke beschränkt, was natürlich einer gewissen Schleichwerbung gleichkommt.</i> Im Sinne des Themas und zur Hinführung auf die Problemstellung, kann man dies aber vertreten.“ (KÜHPAT_A5_MZP1) AB2: „Für den Versuch hatten die Schüler in Einzelarbeit grüne Smarties verwendet. Damit ist der Bezug zur Problemstellung gegeben, <i>aber der Vergleich von verschiedenen Farben war nicht möglich.</i>“ (SIMMON_C6_MZP2) Differenzierung: AB1: Das Ankerbeispiel beginnt mit einer positiven Bewertung, gefolgt von einer Rechtfertigung der Bewertung. Im Anschluss wird abgewogen, also zwei Nachteile der getroffenen Entscheidungen aufgezählt (Kodierung: zweimal Abwägen). Der letzte Satz unterstützt erneut die erste Aussage (Kodierung: Rechtfertigung einer Bewertung). AB2: Der erste Satz ist rein deskriptiv (Kategorie: Beschreibung). Die positive Bewertung folgt durch den Bezug zur Problemstellung, was im Anschluss vom fehlenden Vergleich der verschiedenen Farben relativiert, sprich abgewogen wird. <i>3. Kodierregeln</i> Allgemein: KR1: Pro inhaltlich differenziertem Argument gegen zuvor beschriebene Rechtfertigungen einer Bewertung wird mit Abwägen der Bewertung kodiert. KR2: Wird nach dem Abwägen wieder die erste Rechtfertigung der Bewertung unterstützt, wird mit Rechtfertigung der Bewertung kodiert. Differenzierung: Beispiel: <table border="0"> <tr> <td>Bewertung</td> <td>→ Kodierung: Bewertung (positiv/negativ)</td> </tr> <tr> <td>Nennung Vorteil</td> <td>→ Kodierung: Rechtfertigung der Bewertung</td> </tr> <tr> <td>Nennung Nachteil</td> <td>→ Kodierung: Abwägen</td> </tr> <tr> <td>Argument für Verhalten</td> <td>→ Kodierung: Rechtfertigung der Bewertung</td> </tr> </table>	Bewertung	→ Kodierung: Bewertung (positiv/negativ)	Nennung Vorteil	→ Kodierung: Rechtfertigung der Bewertung	Nennung Nachteil	→ Kodierung: Abwägen	Argument für Verhalten	→ Kodierung: Rechtfertigung der Bewertung
Bewertung	→ Kodierung: Bewertung (positiv/negativ)								
Nennung Vorteil	→ Kodierung: Rechtfertigung der Bewertung								
Nennung Nachteil	→ Kodierung: Abwägen								
Argument für Verhalten	→ Kodierung: Rechtfertigung der Bewertung								
	Zielformulierung der Bewertung <i>1. Definition</i> Allgemein: Reine Zielformulierung Differenzierung: Formulierung eines Ziels, welches für die Unterrichtsstunde geplant war und im Kontext bewertet wird. Hypothesen, die am Anfang des Satzes formuliert werden. <i>2. Ankerbeispiele</i> Allgemein:								

	AB1: „Zu Beginn der Stunde war es mein Ziel, alle Schüler durch den Alltagsbezug zu Cola im Supermarkt zu motivieren. Das bekannte Erfrischungsgetränk eignet sich dafür sehr gut, da es sprichwörtlich in aller Munde ist.“ (KÜHPAT_A5_MZP2) AB2: „Durch den Einstieg in meine Unterrichtsstunde habe ich mein Ziel, die Schüler zu aktivieren, erreicht, da sie durch das Bild selbst nachdenken und ihr Vorwissen aktivieren mussten.“ (SÜSSAR_A2_MZP2) Differenzierung: AB1: Es beginnt mit einer klaren Zielformulierung. Der zweite Satz lässt die positive Bewertung erkennen mit Rechtfertigung der Bewertung. AB2: Durch den Einstieg habe ich mein Ziel erreicht → Kodierung: Bewertung positiv ...mein Ziel die Schüler zu aktivieren → Zielformulierung der Bewertung ...da sie durch das Bild selbst nachdenken und ihr Vorwissen aktivieren mussten → Kodierung: einmalige Rechtfertigung einer Bewertung, da inhaltlich „nachdenken“ und „Vorwissen aktivieren“ sehr ähnlich ist. „Durch den Einstieg“ ist keine zweite Rechtfertigung, da diese im letzten Teilsatz inhaltlich kodiert wird. 3. Kodierregeln Allgemein: KR1: Die Formulierungen müssen eine klare Zielformulierung inkludieren. Beispiel: „Mein Ziel war es...“; „Ich wollte erreichen, dass...“ KR2: Es wird nur eine Zielformulierung der Bewertung kodiert, wenn im Kontext eine positive oder negative Bewertung zu erkennen ist. Differenzierung: KR3: Wenn anschließend ein Verbesserungsvorschlag mit Bezug zur Zielformulierung gemacht wird, erfolgt die Kodierung Zielformulierung des Verbesserungsvorschlags. AB3: „Bei meinem vierten Stundenziel, der Nennung von verschiedenen Chromatographie-Methoden, sollten die Schüler/innen besser eingebunden werden.“ → Kodierung: Zielformulierung des Verbesserungsvorschlags KR4: Wenn es keine klare Zielformulierung ist, wird die dem Kontext entsprechende Rechtfertigung kodiert. AB4: Anschließend erfolgte die Überleitung zum Einstieg in das Experimentieren mit Bildern von weiteren Stoffgemischen, also einem verschmutzten Sandstrand. Damit sollte der Alltagsbezug geschaffen werden und Interesse geweckt werden. → Kodierung: Rechtfertigung der Beschreibung
c)	Alternative
Alternative	1. Definition Allgemein: Alternative als zusätzliche Möglichkeit zur Entscheidung zwischen zwei sich ausschließenden Handlungsmöglichkeiten oder Dingen, im Sinne einer Entweder-oder-Entscheidung

	Differenzierung: Beispiele: Die eigentliche Einschätzung der Situation war positiv und es wird trotzdem eine zusätzliche Möglichkeit genannt, eine Situation zu bewältigen. Oder: Es werden mehrere Verbesserungsvorschläge gemacht. 2. <i>Ankerbeispiele</i> Allgemein: AB1: Ein alternativer Einstieg wäre die Vorstellung des Textes „So eine Gemeinschaft“, in welchem ein Schüler durch eine Textnachricht gemobbt wird und man herausfinden will, wer das war. Durch geschickte Hinleitung seitens des Lehrers stellen die Schüler die Hypothese auf, den Farbstoff des verwendeten Stiftes zu untersuchen und werden so auf das Thema der Farbtrennung (Chromatographie) aufmerksam und das Trennprinzip kann mithilfe von Smarties durchgeführt werden. AB2: „Diese könnte durch kleine Aufgaben in Partnerarbeit oder mit der Think-Pair-Share-Methode weiter strukturiert werden.“ (BIESTE_A3_MZP3) Differenzierung: AB2: Der erste Teilsatz ist ein Verbesserungsvorschlag, mit dem zweiten Teilsatz wird eine weitere Möglichkeit der Umsetzung, also eine Alternative, genannt. 3. <i>Kodierregeln</i> KR1: Eine längere Beschreibung einer alternativen Handlung wird nur einmal pro Absatz kodiert. → AB1 KR2: Werden mehrere Verbesserungsvorschläge gemacht, wird ab dem zweiten Vorschlag mit Alternative kodiert. → AB2 KR3: Wenn trotz einer positiven Bewertung eine weitere alternative Möglichkeit der Umsetzung genannt wird, wird dies mit Alternative kodiert.
	Rechtfertigung der Alternative 1. <i>Definition</i> Allgemein: Begründungen, Rechtfertigungen einer Alternative Differenzierung: Rechtfertigung einer Alternative durch Hypothesen, Vermutung, Aufzählung einzelner Vorteile, Aufzählung einzelner Nachteile oder begründet durch eine persönliche Meinung 2. <i>Ankerbeispiele</i> Allgemein: AB1: „Diese könnte durch kleine Aufgaben in Partnerarbeit oder mit der Think-Pair-Share-Methode weiter strukturiert werden. Ein Argument dafür ist außerdem,

	dass SuS die Antworten zum Teil nicht sofort wussten („Was versteht man unter einem Reinstoff?“) und so die Möglichkeit haben zu überlegen oder sich bei Unsicherheiten auszutauschen. Diese Methodenvielfalt ist nach Hilbert Meyer ein weiteres Merkmal guten Unterrichts.“ (BIESTE_A3_MZP3) AB2: „Eine generelle Alternative wäre es, den Schüler das Arbeitsblatt zu geben und anzuweisen, die Experimente selbstständig durchführen zu lassen. Dadurch würde das selbstbestimmte Lernen gefördert.“ (KÜHPAT_A5_MZP2) Differenzierung: AB1: Die Think-Pair-Share-Methode wird zweimalig gerechtfertigt. AB2: Es wird der Vorteil Förderung des selbstbestimmten Lernens genannt. 3. Kodierregeln Allgemein: KR1: Mehrere folgende Rechtfertigungen eines Alternativvorschlags werden einzeln kodiert, solange sie sich inhaltlich unterscheiden. Ansonsten wird lediglich einmal kodiert. → AB1 KR2: Sich bedingende Punkte werden lediglich einmal als Rechtfertigung kodiert. Differenzierung: KR3: Werden mehrere Verbesserungsvorschläge gemacht, wird ab dem zweiten als Alternative kodiert. Wird anschließend eine Rechtfertigung formuliert, muss abgewogen werden, ob diese zum Verbesserungsvorschlag gemacht wird oder zur Alternative. Ist keine eindeutige Wahl möglich, wird Rechtfertigung eines Verbesserungsvorschlags kodiert. AB2: „Die Schüler sollten besser mit eingebunden werden oder es sollten Rückfragen gestellt werden, um sicher zu gehen, dass die Schüler verschiedene Chromatographie-Methoden auch nennen können.“ (BERCOR_B6_MZP1) Erklärung: 1. Teilsatz: Verbesserungsvorschlag 2. Teilsatz: Alternative 3. Teilsatz: Rechtfertigung des Verbesserungsvorschlags, da eine eindeutige Zuordnung zur Alternative nicht möglich ist.
	Abwägen der Alternative 1. Definition Allgemein: Betrachtung der anderen Seite eines Alternativvorschlags mit oder ohne Rechtfertigung. Differenzierung: Nach einem Alternativvorschlag (mit oder ohne Rechtfertigung) werden Vor- bzw. Nachteile des Alternativvorschlags genannt. 2. Ankerbeispiele AB1: „Eine generelle Alternative wäre es, den Schüler das Arbeitsblatt zu geben und anzuweisen die Experimente selbstständig durchführen zu lassen. Dadurch

	würde das selbstbestimmte Lernen gefördert. Die Methode birgt auch Risiken, wie die Gefahr des Umgangs mit heißen Substanzen. Unter angemessener Ansprache vor dem Experiment wäre dies jedoch möglich.“ (KÜHPAT_A5_MZP2) AB2: „Eine Alternative beziehungsweise eine Erweiterung, die sich hierbei noch anbieten würde, ist ein kurzes Informationsarbeitsblatt über Diabetes zu verteilen. Damit könnte man nämlich auf mögliche Einschränkungen von Max eingehen, die durch seine Krankheit bedingt sind. Somit wäre natürlich auch ein fachübergreifender Bezug hergestellt. Dagegen spricht allerdings, dass man als Lernender zu weit vom eigentlichen Thema weggeführt wird und vom roten Faden der Stunde abkommt.“ (KÜHPAT_A5_MZP2) <i>3. Kodierregeln</i> Allgemein: KR1: Pro inhaltlich differenziertem Argument gegen zuvor beschriebene Rechtfertigungen einer Alternative wird mit Abwägen der Alternative kodiert. KR2: Wird nach dem Abwägen wieder die erste Rechtfertigung des Alternativvorschlags unterstützt, wird mit Rechtfertigung zur Alternative kodiert. Differenzierung: Beispiel: Alternative → Kodierung: Alternative Nennung Vorteil → Kodierung: Rechtfertigung der Alternative Nennung Nachteil → Kodierung: Abwägen der Alternative Argument für Alternative → Kodierung: Rechtfertigung der Alternative
	Zielformulierung der Alternative <i>1. Definition</i> Allgemein: Reine hypothetische Zielformulierung Differenzierung: Formulierung eines Ziels, welches mit der formulierten Alternative zu erreichen wäre. <i>2. Ankerbeispiele</i> AB1: „Dadurch, dass die Chromatographie auf zwei verschiedene Gegenstände angewendet wurde, wäre auch das Ziel erreicht, dass das Trennprinzip auf andere Farbstoffgemische übertragen werden kann.“ (SIMMON_C6_MZP3) <i>3. Kodierregeln</i> Allgemein: KR1: Die Formulierungen müssen eine klare Zielformulierung inkludieren. Beispiel: „Mein Ziel war es...“; „Ich wollte erreichen, dass...“ KR2: Es wird nur Zielformulierung der Alternative kodiert, wenn im Kontext ein Alternativvorschlag erbracht wurde.

	Differenzierung: KR3: Wenn anschließend ein Verbesserungsvorschlag mit Bezug zur Zielformulierung gemacht wird, erfolgt die Kodierung Zielformulierung des Verbesserungsvorschlags. AB3: „Bei meinem vierten Stundenziel, der Nennung von verschiedenen Chromatographie-Methoden, sollten die Schüler/innen besser eingebunden werden.“ → Kodierung: Zielformulierung des Verbesserungsvorschlags

d)	Konsequenz/Verbesserungsvorschlag
Konsequenz/Verbesserungsvorschlag	1. Definition Allgemein: Es wird eine, oft durch Unzufriedenheit mit vergangenen Handlungen, zwingende, mindestens jedoch mögliche Folgerung oder Verbesserung aufgeführt. Differenzierung: Diese Kategorie beinhaltet sowohl Verbesserungsvorschläge zur gehaltenen Stunde als auch etwaige Konsequenzen, welche aus getroffenen Entscheidungen oder Situationen gezogen werden.
	2. Ankerbeispiele Allgemein: AB1: „Ein klarer Arbeitsauftrag ist also unbedingt notwendig, um dies zu erreichen.“ (BIESTE_A3_MZP2) AB2: „Allerdings wären im Umgang mit der Heizplatte und dem Rührfisch Details hilfreich gewesen. Dazu hätte man eine PowerPoint-Folie erstellen können.“ (KÜHPAT_A5_MZP1) AB3: „Die Schüler/innen haben das Experiment recht gut gemacht, ich hätte jedoch mehr auf das Experiment und die Ergebnisse der Schüler/innen eingehen sollen. So hätte ich ein Papier-Chromatogramm der Schüler/innen der ganzen Klasse zeigen können, bei dem ersichtlich ist, wie sich die Farbe Grün in die zwei Farben Gelb und Blau auftrennen. Oder auch, wie die Farbe Braun durch die Papierchromatographie aufgetrennt wird.“ (BERCOR_B6_MZP2) Differenzierung: AB1: Aus dem Kontext erschließt sich, dass die Probandin eine Konsequenz für ihr weiteres Unterrichten gezogen hat. AB2: Hier wird <u>ein</u> Verbesserungsvorschlag kodiert, da die Umsetzung der geforderten Details des ersten Satzes im zweiten Satz nur genauer ausgeführt wird. AB3: <u>Ein</u> Verbesserungsvorschlag, welcher sehr genau beschrieben wird. AB4: Der Protokollierungsvorschlag ist ein weiterer Verbesserungsvorschlag und keine Alternative, da inhaltlich unterscheidbar und zeitlich in der Stunde unterscheidbar.

	3. <i>Kodierregeln</i> Allgemein: KR1: Bei einem ausführlichen Verbesserungsvorschlag mit mehreren Schritten nur einfach kodieren. KR2: Bei mehreren inhaltlich verschiedenen Verbesserungsvorschlägen jeden einzelnen kodieren. Differenzierung: KR3: Erfolgen mehrere verschiedenen Verbesserungsvorschläge zum gleichen Punkt, wird nur der erste mit Verbesserungsvorschlag kodiert. Alle folgenden Vorschläge werden als Alternative kodiert. AB5: „Da die Schüler nicht mitschreiben müssen, wäre es besser gewesen, wenn ich den Begriff der Chromatographie mithilfe einer Folie oder der PowerPoint-Präsentation erklären lasse.“ (BERCOR_B6_MZP1)
	Rechtfertigung der Konsequenz/des Verbesserungsvorschlags 1. <i>Definition</i> Allgemein: Begründungen, Rechtfertigungen eines Verbesserungsvorschlags Differenzierung: Rechtfertigung eines Verbesserungsvorschlags durch Hypothesen, Vermutung, Aufzählung einzelner Vorteile, Aufzählung einzelner Nachteile oder begründet durch eine persönliche Meinung 2. <i>Ankerbeispiele</i> Allgemein: AB1: „Da die Schüler nicht mitschreiben müssen, wäre es besser gewesen, wenn ich den Begriff der Chromatographie mit Hilfe einer Folie oder der PowerPoint-Präsentation erklären lasse.“ (BERCOR_B6_MZP1) AB2: „Die Schüler sollten besser mit eingebunden werden oder es sollten Rückfragen gestellt werden, um sicher zu gehen, dass die Schüler verschiedene Chromatographie-Methoden auch nennen können.“ (BERCOR_B6_MZP1) Differenzierung: AB1: Der Satz beginnt mit der Rechtfertigung des im Anschluss folgenden Verbesserungsvorschlags. 3. <i>Kodierregeln</i> Allgemein: KR1: Mehrere folgende Rechtfertigungen eines Verbesserungsvorschlags werden einzeln kodiert, solange sie sich inhaltlich unterscheiden. Ansonsten wird lediglich einmal kodiert. KR2: Sich bedingende Punkte werden lediglich einmal als Rechtfertigung kodiert.

	Differenzierung: KR3: Werden mehrere Verbesserungsvorschläge gemacht, wird ab dem zweiten als Alternative kodiert. Wird anschließend eine Rechtfertigung formuliert, muss abgewogen werden, ob diese dem Verbesserungsvorschlag zuzuordnen ist oder der Alternative. Ist keine eindeutige Wahl möglich, wird immer Rechtfertigung eines Verbesserungsvorschlags kodiert.
	Abwägen der Konsequenz/des Verbesserungsvorschlags <i>1. Definition</i> Allgemein: Betrachtung der anderen Seite eines Verbesserungsvorschlags Differenzierung: Nach einem Verbesserungsvorschlag (mit oder ohne Rechtfertigung) werden Vor- bzw. Nachteile des Verbesserungsvorschlags genannt. <i>2. Ankerbeispiele</i> Allgemein: AB1: „Dies wäre wichtig gewesen, um mein drittes Stundenziel zu erfüllen, wobei die Schüler/innen ihr Papier-Chromatogramm auswerten sollen. Zwar haben einige Schüler bei ihrem Versuch gesehen, dass sich die Farbe Grün in zwei Farben auf-trennt, ich hätte es jedoch allen Schülern/innen noch einmal deutlich zeigen sollen. Die Schüler/innen konnten aber schon meist zwischen mobiler und stationärer Phase unterscheiden.“ (BERCOR_B6_MZP2) Differenzierung: AB1: Das Beispiel beginnt mit einer Zielformulierung. Durch das Wort „zwar“ wird das Abwägen des im nächsten Teilsatz folgenden Verbesserungsvorschlag deutlich. Im letzten Satz wird die Verbesserung wieder relativiert, was ebenfalls mit Abwägen kodiert wird. <i>3. Kodierregeln</i> Allgemein: KR1: Pro inhaltlich differenziertem Argument gegen zuvor beschriebene Rechtfertigungen eines Verbesserungsvorschlags wird mit Abwägen des Verbesserungsvorschlags kodiert. KR2: Wird nach dem Abwägen wieder die erste Rechtfertigung des Verbesserungsvorschlags unterstützt, wird mit Rechtfertigung des Verbesserungsvorschlags kodiert. Differenzierung: Beispiel: Verbesserungsvorschlag → Kodierung: Verbesserungsvorschlag Nennung Vorteil → Kodierung: Rechtfertigung des Verbesserungsvorschlags Nennung Nachteil → Kodierung: Abwägen des Verbesserungsvorschlags

	Argument für Verbesserungsvorschlag → Kodierung: Rechtfertigung des Verbesserungsvorschlags	
	Zielformulierung der Konsequenz/des Verbesserungsvorschlags	
	1. <i>Definition</i> Allgemein: Reine hypothetische Zielformulierung, die mithilfe eines Verbesserungsvorschlags zu erreichen ist Differenzierung: Formulierung eines möglichen Ziels, welches mit dem ausformulierten Verbesserungsvorschlag zu erreichen ist. 2. <i>Ankerbeispiele</i> AB1: „Da die Schüler nicht mitschreiben müssen, wäre es besser gewesen, wenn ich den Begriff der Chromatographie mit Hilfe einer Folie oder der PowerPoint-Präsentation erklären lasse. Das erste Ziel wäre gewesen, dass die Schüler den Begriff Chromatographie erklären.“ (BERCOR_B6_MZP1) AB2: „Dies wäre wichtig gewesen, um mein drittes Stundenziel zu erfüllen, wobei die Schüler/innen ihr Papier-Chromatogramm auswerten sollen. Zwar haben einige Schüler bei ihrem Versuch gesehen, dass sich die Farbe Grün in zwei Farben auf trennt, ich hätte es jedoch allen Schülern/innen noch einmal deutlich zeigen sollen. Die Schüler/innen konnten aber schon meist zwischen mobiler und stationärer Phase unterscheiden.“ (BERCOR_B6_MZP2) 3. <i>Kodierregeln</i> Allgemein: KR1: Die Formulierungen müssen eine klare Zielformulierung inkludieren. Beispiel: „Mein Ziel war es...“; „Ich wollte erreichen, dass...“ KR2: Es wird nur Zielformulierung des Verbesserungsvorschlags kodiert, wenn im Kontext ein Verbesserungsvorschlag erbracht wurde.	

3. Durchgang des Materials: Kodierung der Reflexionstiefe II

e)	Perspektiven bei Abwägen und Rechtfertigungen	
Perspektivenüber-	Schülerperspektive	
	1. <i>Definition</i> Allgemein: Die Unterrichtssituationen werden aus der Perspektive der Schüler betrachtet. Die Probanden versetzen sich dabei in die Rolle der Schüler und antizipieren deren Gefühle oder Gedanken. Differenzierung:	

	Hineinversetzen in die Gefühlswelt der SuS der vergangenen Unterrichtsstunde. Die Probanden versetzen sich in die Schülerperspektive und machen sich Gedanken über die Motivation und die Bedürfnisse der SuS. Dies wird meist in Rechtfertigungen oder in Abwägungen der Beurteilungen deutlich. Auch Rechtfertigungen und Abwägungen von Alternativen oder Verbesserungsvorschlägen können hypothetisch die Gefühlswelt der SuS einbeziehen. 2. <i>Ankerbeispiele</i> AB1: „Dadurch erkennen die SuS nämlich einen Sinn hinter ihren Experimenten und experimentieren nicht einfach stumpf vor sich hin, weil es der Lehrkörper vorschreibt, ohne den Weg bzw. das Ziel zu kennen.“ (BOXALE_A1_MZP1) AB2: „Aus Sicht der Lernenden könnte dies auch eine Erleichterung und Auflockerung dargestellt haben, da sie ohne Zwang und mit weniger Kontrolle handeln durften.“ (KÜHPAT_A5_MZP3) 3. <i>Kodierregeln</i> Allgemein: KR1: Es werden lediglich die Bewertungen, Rechtfertigungen, Zielformulierungen und die Inhalte der vier Punkte zum Abwägen (beschreibend, bewertend, Alternativen und Konsequenzen/Verbesserungsvorschläge) betrachtet, welche die Gefühlswelt der SuS zum Inhalt haben. KR2: Die Schülerperspektive wird kodiert, wenn eine Vermutung zu Gefühlen, Gedanken, Motivationen oder Bedürfnissen von Schülerinnen und Schülern in der Unterrichtsstunde formuliert wird. KR3: In einem Absatz werden inhaltlich zusammengehörige Perspektiven, welche häufiger angesprochen werden, lediglich einmal kodiert. KR4: Inhaltlich unterscheidbare Perspektiven werden einzeln kodiert. Differenzierung: KR5: Rein beobachtbare Situationen werden nicht der Schülerperspektive zugeordnet. Es muss eine Perspektivenübernahme erfolgen, also wirklich der Blickwinkel aus Schülersicht eingenommen werden. KR6: Hypothetische Ziele im Verbesserungsvorschlag oder bei Alternativen, wie Steigerung der Motivation, werden nicht als Schülerperspektive kodiert, sondern als Theorieperspektive.
	Fachperspektive 1. <i>Definition</i> Allgemein: Die Unterrichtssituationen werden aus der Perspektive der Fachwissenschaft Chemie betrachtet. Entscheidungen werden fachlich begründet. Differenzierung:

	Es wird chemisches Fachwissen als Begründung, Rechtfertigung, beim Abwägen, bei den Zielformulierungen oder bewertend angeführt. Dies sind zum Beispiel Erklärungen zu fachlichen Fehlern oder fachliche Rechtfertigungen. 2. <i>Ankerbeispiele</i> Allgemein: AB1: „Das folgende „Experiment“ dient dazu den SuS aufzuzeigen, dass die Dichte eine Stoffeigenschaft ist, welche den direkten Zusammenhang zwischen Masse und Volumen eines Körpers erklärt.“ (BOXALE_A1_MZP2) AB2: „Der Versuch hat dennoch nicht richtig funktioniert, da zwei Gruppen beim Lösen des Salzes nicht richtig umgerührt haben, wodurch sich das Salz nicht ordentlich lösen konnte, bzw. es wurde zu viel Wasser zugegeben, dass die Salzkonzentration in der Abdampfschale zu gering war.“ (SÜSSAR_A2_MZP3) Differenzierung: AB1: Es wird eine Zielformulierung des Experiments gemacht. Dieses Ziel ist fachwissenschaftlicher Natur. AB2: Nach einer negativen Bewertung folgt eine fachwissenschaftliche Begründung. 3. <i>Kodierregeln</i> Allgemein: KR1: Es werden die vier Arten der Rechtfertigungen, Zielformulierungen und die Punkte des Abwägens (beschreibend, bewertend, Alternativen und Konsequenzen/Verbesserungsvorschläge) betrachtet, welche die Fachwissenschaft Chemie zum Inhalt haben. KR2: Die Fachperspektive wird kodiert, wenn eine fachliche Erklärung als Rechtfertigung, abwägend oder als Zielformulierung genutzt wird. Differenzierung: KR3: Die fachliche Richtigkeit der Perspektivenübernahme ist für die Kodierung unerheblich.

	Theorieperspektive
	<i>1. Definition</i> Allgemein: Der Unterricht wird aus einer chemiedidaktischen oder pädagogischen Perspektive betrachtet. Dabei fließt theoretisches Wissen, welches sowohl auf Erfahrung als auch auf Theorie basieren kann, in die Reflexion mit ein. Differenzierung: Bewertungen, Rechtfertigungen, Abwägungen oder Zielformulierungen werden auf chemiedidaktische oder pädagogische Inhalte bezogen. Dies erfolgt häufig durch Literaturbezüge oder der Beschreibung von möglichen Auswirkungen didaktischer Methoden. Ebenso können beispielsweise Rechtfertigungen von Verbesserungsvorschlägen durch didaktische Konzepte oder allgemeingültige didaktische Aussagen kodiert werden. <i>2. Ankerbeispiele</i> Allgemein: AB1: „Nach Hilbert Meyer zeichnet guten Unterricht aber unter anderem auch eine klare Strukturierung aus, die sich durch einen für alle erkennbaren roten Faden kennzeichnet.“ (BIESTE_A3_MZP3) AB2: „Dieses Vorgehen wäre ganz im Sinne des „demokratischen Unterrichtens.“ (BOXALE_A1_MZP2) Differenzierung: AB1: Als Rechtfertigung für die Einhaltung des roten Fadens wird ein Literaturbezug hergestellt. AB2: Als Rechtfertigung wird ein didaktisches Konzept angeführt. <i>3. Kodierregeln</i> Allgemein: KR1: Es werden die vier Arten der Rechtfertigungen, Zielformulierungen und die Punkte des Abwägens (beschreibend, bewertend, Alternativen und Konsequenzen/Verbesserungsvorschläge) betrachtet, welche Themen der universitären Ausbildung der Fachdidaktik Chemie zum Inhalt haben. KR2: Die Theorieperspektive wird kodiert, wenn eine didaktische oder pädagogische Erklärung als Rechtfertigung, abwägend oder als Zielformulierung genutzt wird. Differenzierung: KR3: Didaktische oder pädagogische Erklärungen beinhalten auch die Hypothese der Steigerung der Motivation der SuS bei Verbesserungsvorschlägen und Alternativen. KR4: Das Einhalten von Sicherheitsregeln wird nicht als Fachperspektive kodiert.

	Politisch/Soziale Perspektive

	<i>1. Definition</i> Allgemein: Der Unterricht wird aus einer politischen oder sozialen Perspektive betrachtet. Es wird über bestehende gesellschaftliche Systeme nachgedacht. Differenzierung: Bewertungen, Rechtfertigungen, Abwägungen oder Zielformulierungen werden auf soziale Strukturen oder politische Inhalte bezogen. Die erfolgt beispielsweise durch politisch/soziale Überlegungen über das Bildungssystem. <i>2. Ankerbeispiele</i> AB1: „Denkanstöße wie Qualitätssicherung in Lebensmitteln, Kontrolle der Inhaltsstoffe und Inhaltsangaben auf Lebensmittelverpackungen hätten den Chemieunterricht direkt in einen soziopolitischen Kontext setzen können, welcher bei den Schülern ein neues chemisches Bewusstsein im Alltag hätte prägen können.“ (BAD-CHR_C6_MZP3) AB2: „Hier lässt sich starke Kritik am gängigen Schulsystem einbringen, da die Schüler selbst durch ausreichend Pausen wenig Konzentration aufbringen, wenn sie so lange geistig gefordert sind. Hier sollte man sich vom Ministerium oder auch schulintern auf ein System einigen können, dass die Schüler individuell in ihren Leistungsfähigkeiten auffängt und so die Konzentration und Motivation in den Stunden fördern kann.“ (SEIPAT_A6_MZP1) <i>3. Kodierregeln</i> Allgemein: KR1: Es werden die vier Arten der Rechtfertigungen, Zielformulierungen und die Punkte des Abwägens (beschreibend, bewertend, Alternativen und Konsequenzen/Verbesserungsvorschläge) betrachtet, welche politisch/soziale Themen zum Inhalt haben. KR2: Die politisch/soziale Perspektive wird kodiert, wenn eine gesellschaftspolitische oder soziale Erklärung als Rechtfertigung, abwägend oder als Zielformulierung genutzt wird. Differenzierung: KR3: Die politische Richtung oder Korrektheit ist für die Kodierung unerheblich.

f)	Literaturbezug
	<i>1. Definition</i> Allgemein: Es werden direkte oder indirekte Zitate durch Literaturangaben belegt. Differenzierung: Als Literaturquellen gelten alle Quellenangaben, ob wissenschaftlich oder nichtwissenschaftlich. So wird auch ein Lehrplanbezug als Literaturquelle kodiert.

2. Ankerbeispiele

Allgemein:

AB1: „Laut Hilbert Meyer kann man diese vor allem durch einen Wechsel der Sozialformen erreichen.“ (KÜHPAT_A5_MZP3)

AB2: „Andererseits wollte ich die Schüler aber frei arbeiten lassen, da sie den naturwissenschaftlichen Erkenntnisgewinn eigenständig durchleben sollten, da dieser im Lehrplan als sehr wichtig hervorgehoben wird.“ (SÜSSAR_A2_MZP3)

Differenzierung:

AB2: Der Lehrplan wird als Begründung herangezogen.

3. Kodierregeln

Allgemein:

KR1: Wird eine Quelle genannt, so wird diese mit Literaturbezug kodiert.

KR2: Inhaltlich gleiche Aussagen zum gleichen Punkt werden lediglich einfach kodiert.

KR3: Zur Begründung inhaltlich verschiedener Punkte kann die gleiche Literaturquelle mehrfach kodiert werden.

Differenzierung:

KR4: Der Bezug zu Lehrplänen wird ebenso als Literaturquelle kodiert.

KR5: Die Korrektheit der Quellenangabe ist für die Kodierung unerheblich.

10.3. Statistiken

10.3.1. Statistiken zum Kapitel 5.1

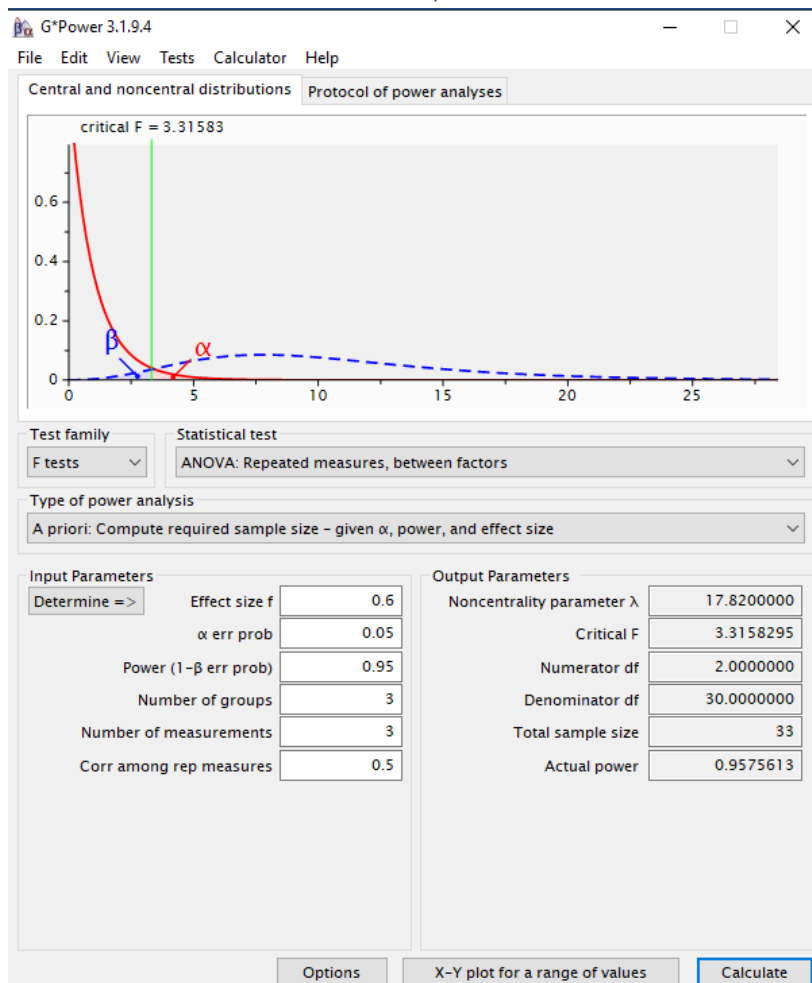


Abbildung 87: Berechnung des Stichprobenumfangs mithilfe der Software G*Power

10.3.2. Statistiken zum Kapitel 5.4.1

Tabelle 174: Berechnungen zur Kriteriumsvalidität des Reflexionskompetenztests

Korrelationen			Scores Refl.- Post	Scores CK_Post
Spearman-Rho	Scores Refl.-Post	Korrelationskoeffizient	1.000	.319
		Sig. (2-seitig)	.	.047
		N	39	39
	Scores CK_Post	Korrelationskoeffizient	.319	1.000
		Sig. (2-seitig)	.047	.
		N	39	39

*. Die Korrelation ist auf dem 0.05 Niveau signifikant (zweiseitig).

Berechnungen zur korrelativen Gültigkeit des Kodiermanuals

Tabelle 175: Berechnungen zur korrelativen Gültigkeit des Kodiermanuals

Korrelationen			SummexTiefe	xLernen von Chemie in der Schule
Spearman-Rho	SummexTiefe	Korrelationskoeffizient	1.000	-.265

xLernen von Chemie in der Schule	Sig. (2-seitig)	.	.045
	N	58	58
	Korrelationskoeffizient	-.265 [*]	1.000
	Sig. (2-seitig)	.045	.
	N	58	58

*. Die Korrelation ist auf dem 0.05 Niveau signifikant (zweiseitig).

Korrelationen

			SummerxPer- spektiven	xBestrebungen im Stu- dium	xLernen von Chemie in der Schule
Spearman-Rho	SummerxPerspektiven	Korrelationskoeffizient	1.000	.304 [*]	-.279 [*]
		Sig. (2-seitig)	.	.020	.034
		N	58	58	58
	xBestrebungen im Stu- dium	Korrelationskoeffizient	.304 [*]	1.000	-.099
		Sig. (2-seitig)	.020	.	.460
		N	58	58	58
	xLernen von Chemie in der Schule	Korrelationskoeffizient	-.279 [*]	-.099	1.000
		Sig. (2-seitig)	.034	.460	.
		N	58	58	58

*. Die Korrelation ist auf dem 0.05 Niveau signifikant (zweiseitig).

10.3.3. Statistiken zum Kapitel 6.1.1

Tabelle 176: Reliabilitätsanalyse – PCK-Test – Pilot- und Hauptstudie

Reliabilitätsstatistiken	
Cronbachs Alpha	Anzahl der Items
.854	38

	Item-Skala-Statistiken			
	Skalenmittelwert, wenn I- tem weggelassen	Skalenvarianz, wenn I- tem weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
yPCK_Item1_12	26.1754	37.504	.266	.852
yPCK_Item1_13	26.4035	36.602	.336	.850
yPCK_Item1_24	26.1053	37.882	.249	.852
yPCK_Item1_34	26.3509	36.696	.333	.850
yPCK_Item2_12	26.9123	37.724	.268	.852
yPCK_Item2_24	26.1404	37.587	.281	.851
yPCK_Item2_34	26.1579	36.992	.404	.849
yPCK_Item3_13	26.4386	36.429	.360	.850
yPCK_Item3_23	26.1228	37.003	.462	.848
yPCK_Item3_24	26.4211	36.355	.375	.849
yPCK_Item3_34	26.1579	37.314	.328	.850
yPCK_Item4_12	26.2807	37.027	.299	.851
yPCK_Item4_23	26.5263	36.504	.341	.850
yPCK_Item4_24	26.1228	37.967	.203	.853
yPCK_Item5_12	26.4211	36.284	.387	.849
yPCK_Item5_13	26.1754	36.540	.486	.847
yPCK_Item5_24	26.2982	36.642	.363	.850
yPCK_Item5_34	26.1754	36.897	.404	.849
yPCK_Item7_12	26.2807	36.777	.346	.850
yPCK_Item7_23	26.1404	36.944	.443	.848
yPCK_Item7_24	26.2456	36.403	.442	.848
yPCK_Item8_12	26.2281	36.393	.459	.847
yPCK_Item8_14	26.2281	36.643	.407	.849
yPCK_Item10_12	26.3333	36.869	.307	.851
yPCK_Item10_14	26.4211	36.534	.344	.850
yPCK_Item11_14	26.2807	36.491	.401	.849
yPCK_Item11_34	26.5614	36.786	.295	.852
yPCK_Item12_12	26.2807	36.563	.387	.849
yPCK_Item12_13	26.3333	36.619	.352	.850
yPCK_Item13_12	26.3509	36.946	.288	.852
yPCK_Item13_34	26.1579	37.171	.362	.850
yPCK_Item14_12	26.0526	38.015	.345	.851
yPCK_Item14_14	26.0702	37.959	.299	.852
yPCK_Item15_12	26.1053	37.810	.270	.852
yPCK_Item15_13	26.4912	37.326	.205	.854
yPCK_Item17_13	26.5965	36.638	.324	.851
yPCK_Item17_23	26.3158	36.506	.380	.849
yPCK_Item18_12	26.7895	37.419	.240	.853

10. Anhang

10.3.4. Statistiken zum Kapitel 6.1.2

Tabelle 177: Reliabilitätsanalyse – CK-Test – Pilotstudie

Reliabilitätsstatistiken					
Cronbachs Alpha		Anzahl der Items			
.727		19			
Item-Skala-Statistiken					
	Skalenmittelwert, wenn I-tem weggelassen	Skalenvarianz, wenn I-tem weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen	
yCK_Item1 Punkte	10.3846	11.090	.518	.700	
yCK_Item2 Punkte	10.5385	12.103	.058	.738	
yCK_Item3 Punkte	10.3846	11.256	.448	.705	
yCK_Item4 Punkte	10.3077	12.064	.200	.724	
yCK_Item5 Punkte	10.6154	11.756	.148	.731	
yCK_Item6 Punkte	10.6154	12.090	.051	.740	
yCK_Item8 Punkte	10.5385	11.269	.318	.714	
yCK_Item9 Punkte	11.0769	12.244	.054	.734	
yCK_Item10 Punkte	10.4615	11.436	.303	.716	
yCK_Item11 Punkte	10.7692	10.359	.568	.688	
yCK_Item12 Punkte	10.6923	11.064	.345	.712	
yCK_Item14 Punkte	10.8462	9.808	.776	.666	
yCK_Item15 Punkte	10.7692	11.192	.307	.716	
yCK_Item16 Punkte	10.6923	11.064	.345	.712	
yCK_Item17 Punkte	10.9231	10.744	.493	.697	
yCK_Item18 Punkte	10.3846	12.256	.049	.734	
yCK_Item19 Punkte	11.0000	12.167	.054	.736	
yCK_Item20 Punkte	10.3077	11.897	.288	.719	
yCK_Item21 Punkte	10.8462	10.974	.386	.708	

Tabelle 178: Reliabilitätsanalyse – CK-Test – Pilot- und Hauptstudie

Reliabilitätsstatistiken				
Cronbachs Alpha		Anzahl der Items		
.734		18		
Item-Skala-Statistiken				
	Skalenmittelwert, wenn I-tem weggelassen	Skalenvarianz, wenn I-tem weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
yCK_Item1 Punkte	10.6596	10.577	.419	.714
yCK_Item2 Punkte	10.6596	10.838	.314	.722
yCK_Item3 Punkte	10.6809	10.744	.334	.721
yCK_Item4 Punkte	10.5957	10.898	.363	.720
yCK_Item5 Punkte	10.8511	10.738	.261	.727
yCK_Item6 Punkte	10.7872	11.475	.038	.747
yCK_Item8 Punkte	10.6596	10.751	.349	.720
yCK_Item9 Punkte	11.0213	10.717	.258	.728
yCK_Item10 Punkte	10.6596	10.838	.314	.722
yCK_Item11 Punkte	10.9787	9.934	.512	.702
yCK_Item12 Punkte	10.8511	10.956	.191	.734
yCK_Item14 Punkte	11.0638	10.061	.481	.705
yCK_Item15 Punkte	11.1489	11.260	.107	.741
yCK_Item16 Punkte	10.8936	10.749	.250	.729
yCK_Item17 Punkte	11.0638	10.105	.466	.707
yCK_Item18 Punkte	10.7234	10.683	.327	.721
yCK_Item20 Punkte	10.5745	11.032	.333	.723
yCK_Item21 Punkte	11.0851	10.384	.377	.716

10.3.5. Statistiken zum Kapitel 6.1.3

Tabelle 179: Reliabilitätsanalyse – Reflexionskompetenztest – Pilotstudie

Reliabilitätsstatistiken				
Cronbachs Alpha	Anzahl der Items			
.807	12			

Item-Skala-Statistiken				
	Skalenmittelwert, wenn I-tem weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn I-tem weggelassen
yRefl_Item1 Punkte	6.7059	6.721	.465	.793
yRefl_Item3 Punkte	6.8824	6.235	.502	.789
yRefl_Item4 Punkte	7.5294	7.140	.335	.803
yRefl_Item5 Punkte	6.7059	6.846	.389	.799
yRefl_Item8 Punkte	6.7647	6.441	.520	.787
yRefl_Item9 Punkte	6.7059	6.846	.389	.799
yRefl_Item10 Punkte	6.9412	6.184	.492	.790
yRefl_Item11 Punkte	6.9412	5.809	.666	.770

yRefI_Item13 Punkte	6.6471	7.243	.253	.807
yRefI_Item14 Punkte	7.1765	6.154	.485	.791
yRefI_Item19 Punkte	7.0588	5.809	.629	.774
yRefI_Item20 Punkte	7.4118	6.882	.289	.807

10.3.6. Statistiken zum Kapitel 6.1.4

Tabelle 180: Reliabilitätsanalyse – Studierendenfragebogen – Zielorientierungen – Pilotstudie

Reliabilitätsstatistiken					
Cronbachs Alpha	Anzahl der Items				
.665	16	Item-Skala-Statistiken			
	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen	
yII_Item1.	29.4727	26.884	.118	.666	
yII_Item2	28.5273	26.069	.105	.674	
yII_Item3	29.4545	27.141	.074	.669	
yII_Item4	29.4727	25.809	.305	.649	
yII_Item5	29.7455	27.823	-.025	.675	
yII_Item6	29.0909	25.640	.152	.668	
yII_Item8	29.4727	26.402	.201	.658	
yII_Item9	29.3636	26.236	.177	.661	
yII_Item7 rekodiert	28.3636	25.643	.130	.673	
yII_Item10 rekodiert	29.4364	24.325	.418	.633	
yII_Item11 rekodiert	28.6727	22.113	.459	.618	
yII_Item12 rekodiert	28.9636	22.591	.474	.618	
yII_Item13 rekodiert	29.4182	25.914	.251	.653	
yII_Item14 rekodiert	28.9818	23.389	.439	.626	
yII_Item15 rekodiert	27.9273	21.809	.412	.627	
yII_Item16 rekodiert	28.9091	23.418	.509	.619	

Tabelle 181: Reliabilitätsanalyse – Studierendenfragebogen – Fachspezifische Überzeugungen – Subskala Chemie als Wissenschaft – Pilotstudie

Reliabilitätsstatistiken					
Cronbachs Alpha	Anzahl der Items				
.529	10	Item-Skala-Statistiken			
	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen	
yIII_Item1	19.6379	12.621	.077	.539	
yIII_Item3	19.6552	12.160	.163	.519	
yIII_Item5	18.9310	12.451	.033	.562	
yIII_Item8	19.5862	12.212	.146	.523	
yIII_Item2 rekodiert	18.8448	11.291	.229	.501	
yIII_Item4 rekodiert	19.4828	9.377	.529	.390	
yIII_Item6 rekodiert	20.0172	11.877	.214	.506	
yIII_Item7 rekodiert	20.0862	12.431	.140	.523	
yIII_Item9 rekodiert	19.6552	10.756	.305	.476	
yIII_Item10 rekodiert	19.7759	10.142	.394	.443	

Tabelle 182: Reliabilitätsanalyse – Studierendenfragebogen – Fachspezifische Überzeugungen – Subskala Lehrern von Chemie in der Schule – Pilotstudie

Reliabilitätsstatistiken					
Cronbachs Alpha	Anzahl der Items				
.776	21	Item-Skala-Statistiken			
	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen	
yIII_Item11	41.3276	40.329	.417	.764	
yIII_Item12	41.2069	39.044	.460	.759	
yIII_Item16	41.1552	38.975	.517	.757	
yIII_Item17	40.9655	40.209	.322	.768	
yIII_Item18	41.3793	39.889	.494	.760	
yIII_Item19	41.1724	40.005	.365	.765	
yIII_Item20	40.6207	42.380	.064	.784	
yIII_Item22	41.2414	41.660	.260	.771	
yIII_Item23	40.5000	43.377	-.031	.789	
yIII_Item24	40.5345	41.867	.163	.777	
yIII_Item25	40.4655	39.376	.329	.768	
yIII_Item26	41.0172	41.280	.243	.772	
yIII_Item27	40.0172	38.403	.420	.761	
yIII_Item28	40.8793	38.283	.588	.752	
yIII_Item13 rekodiert	40.2759	39.361	.360	.765	
yIII_Item14 rekodiert	41.0345	40.034	.425	.763	
yIII_Item15 rekodiert	39.8621	39.174	.363	.765	

10. Anhang

yIII_Item21 rekodiert	40.1207	39.406	.443	.761
yIII_Item29 rekodiert	40.2759	39.010	.383	.764
yIII_Item30 rekodiert	40.1724	39.900	.210	.779
yIII_Item31 rekodiert	40.2586	37.949	.381	.764

Tabelle 183: Reliabilitätsanalyse – Studierendenfragebogen – Fachspezifische Überzeugungen – Subskala Lernen von Chemie in der Schule – Pilotstudie

Reliabilitätsstatistiken				
Cronbachs Alpha	Anzahl der Items			
.499	8			
Item-Skala-Statistiken				
	Skalenmittelwert, wenn I-tem weggelassen	Skalenvarianz, wenn I-tem weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
yIII_Item35	16.5172	6.815	.337	.439
yIII_Item36	15.1897	6.683	.141	.504
yIII_Item38	16.4310	6.109	.542	.367
yIII_Item39	15.1034	8.375	-.220	.653
yIII_Item32 rekodiert	15.7241	5.256	.594	.299
yIII_Item33 rekodiert	16.0345	7.051	.218	.471
yIII_Item34 rekodiert	15.5000	5.798	.331	.417
yIII_Item37 rekodiert	15.6207	6.766	.202	.474

Tabelle 184: Reliabilitätsanalyse – Studierendenfragebogen – Fachspezifische Überzeugungen – Gesamt

Reliabilitätsstatistiken					
Cronbachs Alpha		Anzahl der Items			
.816		39			
Item-Skala-Statistiken					
	Skalenmittelwert, wenn I-tem weggelassen	Skalenvarianz, wenn I-tem weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen	
yIII_Item1	80.3793	110.099	.127	.817	
yIII_Item3	80.3966	114.068	-.132	.825	
yIII_Item5	79.6724	112.926	-.063	.826	
yIII_Item8	80.3276	110.750	.078	.819	
yIII_Item11	81.0862	107.799	.380	.811	
yIII_Item12	80.9655	105.999	.415	.809	
yIII_Item16	80.9138	105.378	.503	.807	
yIII_Item17	80.7241	106.449	.382	.810	
yIII_Item18	81.1379	107.595	.408	.810	
yIII_Item19	80.9310	108.627	.243	.814	
yIII_Item20	80.3793	110.099	.115	.818	
yIII_Item22	81.0000	109.614	.253	.814	
yIII_Item23	80.2586	111.985	.001	.821	
yIII_Item24	80.2931	108.737	.257	.813	
yIII_Item25	80.2241	106.493	.304	.812	
yIII_Item26	80.7759	108.738	.263	.813	
yIII_Item27	79.7759	106.317	.310	.812	
yIII_Item28	80.6379	103.428	.635	.803	
yIII_Item35	80.9828	106.368	.530	.808	
yIII_Item36	79.6552	109.318	.137	.818	
yIII_Item38	80.8966	104.586	.626	.804	
yIII_Item39	79.5690	118.039	-.317	.834	
yIII_Item2 rekodiert	79.5862	103.826	.424	.808	
yIII_Item4 rekodiert	80.2241	101.616	.502	.804	
yIII_Item6 rekodiert	80.7586	105.835	.401	.809	
yIII_Item7 rekodiert	80.8276	111.128	.069	.818	
yIII_Item9 rekodiert	80.3966	103.717	.413	.808	
yIII_Item10 rekodiert	80.5172	105.623	.298	.812	
yIII_Item13 rekodiert	80.0345	107.192	.285	.813	
yIII_Item14 rekodiert	80.7931	106.167	.487	.808	
yIII_Item15 rekodiert	79.6207	107.047	.281	.813	
yIII_Item21 rekodiert	79.8793	106.178	.425	.809	
yIII_Item29 rekodiert	80.0345	106.841	.296	.812	
yIII_Item30 rekodiert	79.9310	105.504	.289	.813	
yIII_Item31 rekodiert	80.0172	102.263	.461	.806	
yIII_Item32 rekodiert	80.1897	102.963	.554	.804	
yIII_Item33 rekodiert	80.5000	108.009	.350	.811	
yIII_Item34 rekodiert	79.9655	102.139	.525	.804	
yIII_Item37 rekodiert	80.0862	107.343	.311	.812	

Tabelle 185: Reliabilitätsanalyse – Selbstwirksamkeitserwartung – Subskala Planung von Experimenten – Pilotstudie

Reliabilitätsstatistiken				
Cronbachs Alpha	Anzahl der Items			
.665	5			
Item-Skala-Statistiken				
	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn I-tem weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
yIV_Item1	10.70	4.213	.525	.562
yIV_Item2	10.42	4.712	.364	.637

yIV_Item3	10.51	4.862	.365	.636
yIV_Item4	9.96	4.534	.568	.560
yIV_Item5	10.82	4.183	.338	.669

Tabelle 186: Reliabilitätsanalyse – Selbstwirksamkeitserwartung – Subskala Durchführung von Experimenten – Pilotstudie

Reliabilitätsstatistiken				
Cronbachs Alpha		Anzahl der Items		
.735		4		
Item-Skala-Statistiken				
	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
yIV_Item6	7.58	2.712	.520	.680
yIV_Item7	7.40	2.888	.416	.734
yIV_Item8	7.39	2.170	.635	.606
yIV_Item9	6.95	2.551	.547	.664

Tabelle 187: Reliabilitätsanalyse – Selbstwirksamkeitserwartung – Subskala Didaktische Rekonstruktion unterrichts-relevanter Inhalte – Pilotstudie

Reliabilitätsstatistiken				
Cronbachs Alpha	Anzahl der Items			
.767	8			
Item-Skala-Statistiken				
	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
yIV_Item10	17.40	9.752	.455	.744
yIV_Item11	17.31	9.551	.458	.743
yIV_Item12	17.50	10.184	.374	.756
yIV_Item13	17.19	8.542	.566	.723
yIV_Item14	17.52	8.710	.705	.701
yIV_Item15	17.83	10.250	.258	.777
yIV_Item16	17.36	9.042	.511	.734
yIV_Item17	17.05	9.489	.426	.749

Tabelle 188: Reliabilitätsanalyse – Selbstwirksamkeitserwartung – Subskala Didaktische Rekonstruktion im Unterricht – Pilotstudie

Reliabilitätsstatistiken					
Cronbachs Alpha		Anzahl der Items			
.685		5			
Item-Skala-Statistiken					
	Skalenmittelwert, wenn I- tem weggelassen	Skalenvarianz, wenn I- tem weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen	
yIV_Item18	9.21	4.132	.352	.669	
yIV_Item19	9.48	4.008	.330	.678	
yIV_Item20	9.29	3.369	.462	.625	
yIV_Item21	9.45	3.339	.586	.568	
yIV_Item22	9.33	3.347	.478	.617	

Tabelle 189: Reliabilitätsanalyse – Selbstwirksamkeitserwartung – Gesamt – Pilotstudie

Reliabilitätsstatistiken					
Cronbachs Alpha		Anzahl der Items			
.876		22			
Item-Skala-Statistiken					
	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen	
yIV_Item1	52.12	62.003	.497	.870	
yIV_Item2	51.84	62.778	.443	.872	
yIV_Item3	51.93	64.852	.294	.876	
yIV_Item4	51.39	63.741	.444	.872	
yIV_Item5	52.25	60.081	.505	.871	
yIV_Item6	52.33	62.726	.548	.869	
yIV_Item7	52.16	63.278	.487	.871	
yIV_Item8	52.14	60.766	.607	.866	
yIV_Item9	51.70	63.213	.459	.871	
yIV_Item10	52.04	62.927	.536	.869	
yIV_Item11	51.95	63.908	.397	.873	
yIV_Item12	52.16	65.457	.304	.876	
yIV_Item13	51.82	59.933	.633	.865	
yIV_Item14	52.18	62.504	.543	.869	
yIV_Item15	52.47	64.754	.296	.876	
yIV_Item16	51.98	62.768	.463	.871	
yIV_Item17	51.68	62.470	.495	.870	
yIV_Item18	52.05	65.015	.363	.874	
yIV_Item19	52.32	65.863	.228	.878	
yIV_Item20	52.11	62.310	.487	.871	
yIV_Item21	52.28	62.098	.563	.868	
yIV_Item22	52.16	61.421	.552	.868	

10.3.7. Statistiken zum Kapitel 6.1.5

Tabelle 190: Kodiermanual – Intracoderreliabilität – Pilotstudie

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      LL95%CI      UL95%CI      Units      Observrs      Pairs
Nominal      ,6846      ,6091      ,7600     165,0000      2,0000     165,0000
Probability (q) of failure to achieve an alpha of at least alphamin:
      alphamin      q
      ,9000      1,0000
      ,8000      ,9989
      ,7000      ,6610
      ,6700      ,3239
      ,6000      ,0170
      ,5000      ,0000
Number of bootstrap samples:
      10000
Judges used in these computations:
      Tinal      Tina2
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 191: Intercoderreliabilität – Kodiermanual – Pilotstudie gesamt – Arbeitsschritt 2

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,7392     84,0000     2,0000     84,0000
Judges used in these computations:
      Codierer Codier_1
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 192: Intercoderreliabilität – Kodiermanual – Reflexionsbreite – Pilotstudie

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,7837     16,0000     2,0000     16,0000
Judges used in these computations:
      Codierer Codier_1
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 193: Intercoderreliabilität – Kodiermanual – Reflexionstiefe – Pilotstudie

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,6998     68,0000     2,0000     68,0000
Judges used in these computations:
      Codierer Codier_1
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 194: Intercoderreliabilität – Kodiermanual – Hauptstudie gesamt

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,8847    735,0000     2,0000     735,0000
Judges used in these computations:
      Codierer Codier_1
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 195: Intercoderreliabilität – Kodiermanual – Reflexionsbreite – Hauptstudie

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,8600    137,0000      2,0000    137,0000
Judges used in these computations:
  Codierer Codier_1
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 196: Intercoderreliabilität – Kodiermanual – Reflexionstiefe – Hauptstudie

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,8829    598,0000      2,0000    598,0000
Judges used in these computations:
  Codierer Codier_1
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 197: Intercoderreliabilität – Kodiermanual – Pilotstudie gesamt – Arbeitsschritt 3

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,6138    172,0000      2,0000    172,0000
Judges used in these computations:
  Tina      Linda
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 198: Intercoderreliabilität – Kodiermanual – Reflexionsbreite – Pilotstudie

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,7123     32,0000      2,0000     32,0000
Judges used in these computations:
  Tina      Linda
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 199: Intercoderreliabilität – Kodiermanual – Reflexionstiefe – Pilotstudie

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,5652    140,0000      2,0000    140,0000
Judges used in these computations:
  Tina      Linda
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 200: Intercoderreliabilität – Kodiermanual – Hauptstudie gesamt

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,8791   1070,0000      2,0000   1070,0000
Judges used in these computations:
  Codierer Codier_1
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 201: Intercorderreliabilität – Kodiermanual – Reflexionsbreite – Hauptstudie

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,8631    166,0000      2,0000    166,0000
Judges used in these computations:
  Codierer Codier_1
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

Tabelle 202: Intercorderreliabilität – Kodiermanual – Reflexionstiefe – Hauptstudie

```
Run MATRIX procedure:
Krippendorff's Alpha Reliability Estimate
      Alpha      Units      Obsrvrs      Pairs
Nominal      ,8753    905,0000      2,0000    905,0000
Judges used in these computations:
  Codierer Codier_1
Examine output for SPSS errors and do not interpret if any are found
----- END MATRIX -----
```

10.3.8. Statistiken zum Kapitel 6.2.1

Tabelle 203: Reliabilitätsanalyse – CK-Test – SuS

Reliabilitätsstatistiken	
Cronbachs Alpha	Anzahl der Items
.501	17

	Item-Skala-Statistiken			
	Skalenmittelwert. wenn Item wegge- lassen	Skalenvarianz. wenn Item wegge- lassen	Korrigierte Item- Skala-Korrelation	Cronbachs Alpha. wenn Item wegge- lassen
yCK_Item6 Punkte	8.4079	6.627	.137	.492
yCK_Item7 Punkte	8.3105	6.795	.092	.500
yCK_Item10 Punkte	8.3421	6.521	.203	.478
yCK_Item11 Punkte	8.5579	6.342	.237	.470
yCK_Item12 Punkte	8.4842	6.229	.290	.458
yCK_Item13 Punkte	8.4500	6.264	.282	.460
yCK_Item17 Punkte	8.4263	6.509	.183	.482
yCK_Item18 Punkte	8.7026	6.658	.122	.495
yCK_Item19 Punkte	8.7105	6.692	.110	.497
yCK_Item20 Punkte	8.3842	6.485	.204	.478
yCK_Item21 Punkte	8.7132	6.406	.231	.472
yCK_Item23 Punkte	8.8132	7.039	-.015	.519
yCK_Item25 Punkte	8.7816	6.604	.167	.486
yCK_Item26 Punkte	8.3526	6.614	.158	.487
yCK_Item28 Punkte	8.4868	6.514	.171	.485
yCK_Item29 Punkte	8.6447	6.541	.161	.487
yCK_Item30 Punkte	8.4000	6.832	.055	.509

10.3.9. Statistiken zum Kapitel 6.2.2

Tabelle 204: Reliabilitätsanalyse – Kognitiver Fähigkeitentest – SuS

Reliabilitätsstatistiken	
Cronbachs Alpha	Anzahl der Items
.864	25

	Item-Skala-Statistiken			
	Skalenmittelwert. wenn I- tem weggelassen	Skalenvarianz. wenn I- tem weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha. wenn Item weggelassen
kft_Item1 Punkte	13.1613	29.740	.518	.856
kft_Item2 Punkte	12.5806	31.052	.302	.862
kft_Item3 Punkte	12.6774	31.826	.085	.868
kft_Item4 Punkte	13.0000	29.600	.476	.857
kft_Item5 Punkte	13.1290	29.783	.488	.857
kft_Item6 Punkte	12.6774	29.226	.630	.852
kft_Item7 Punkte	13.0968	30.490	.331	.861
kft_Item8 Punkte	12.9032	28.957	.592	.853
kft_Item9 Punkte	12.6129	30.645	.370	.860
kft_Item10 Punkte	12.9032	30.024	.390	.860
kft_Item11 Punkte	12.9032	29.890	.415	.859
kft_Item12 Punkte	13.0323	28.299	.742	.848
kft_Item13 Punkte	12.8065	29.895	.427	.858

10. Anhang

kft_Item14 Punkte	12.5806	31.185	.270	.863
kft_Item15 Punkte	12.7742	30.314	.355	.861
kft_Item16 Punkte	12.9032	29.224	.541	.855
kft_Item17 Punkte	12.7419	30.598	.310	.862
kft_Item18 Punkte	12.7419	30.198	.388	.860
kft_Item19 Punkte	13.2581	30.198	.514	.857
kft_Item20 Punkte	12.8710	28.449	.694	.849
kft_Item21 Punkte	13.0968	31.424	.151	.867
kft_Item22 Punkte	13.0645	30.729	.276	.863
kft_Item23 Punkte	12.8387	28.406	.709	.849
kft_Item24 Punkte	12.9032	29.957	.402	.859
kft_Item25 Punkte	12.8065	31.828	.068	.870

Tabelle 205: Kognitiver Fähigkeitentest – Normalverteilung

Tests auf Normalverteilung						
Intelligenzquotient	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistik	df	Signifikanz	Statistik	df	Signifikanz
Intelligenzquotient	.205	96	.000	.814	96	.000

a. Signifikanzkorrektur nach Lilliefors

Tabelle 206: Kognitiver Fähigkeitentest – Test auf Unterscheidung der Gruppen – Parametrisch

Test bei unabhängigen Stichproben									
Intelligenz-quotient		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
Intelligenz-quotient	Varianzen sind gleich	.047	.828	-1.073	480	.284	-1.40356	1.30845	-3.97456 1.16743
	Varianzen sind nicht gleich			-1.073	479.502	.284	-1.40356	1.30858	-3.97482 1.16769

Tabelle 207: Kognitiver Fähigkeitentest – Test auf Unterscheidung der Gruppen – Nicht-parametrisch

gruppe Gruppeneinteilung	N	Mittlerer Rang	Rangsumme
Feedback	240	233.94	56144.50
Kein Feedback	242	249.00	60258.50
Gesamt	482		

Statistik für Test^a

Intelligenzquotient	
Mann-Whitney-U	27224.500
Wilcoxon-W	56144.500
Z	-1.191
Asymptotische Signifikanz (2-seitig)	.234

a. Gruppenvariable: gruppe Gruppeneinteilung

10.3.10. Statistiken zum Kapitel 6.2.3

Tabelle 208: Reliabilitätsanalyse – Fachinteresse – Subskala Freude und Interesse an Chemie

Reliabilitätsstatistiken				
Cronbachs Alpha	Anzahl der Items			
.884	5			

Item-Skala-Statistiken				
	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
inter_Item1	12.48	15.770	.759	.852
inter_Item2	11.18	15.478	.651	.875
inter_Item3	11.62	14.991	.687	.867
inter_Item4	12.43	14.883	.748	.852
inter_Item5	12.70	14.881	.772	.847

Tabelle 209: Reliabilitätsanalyse – Fachinteresse – Subskala Chemiebezogenes Fähigkeitsselbstkonzept

Reliabilitätsstatistiken	
Cronbachs Alpha	Anzahl der Items
.842	6

Item-Skala-Statistiken

	Skalenmittelwert, wenn I-tem weggelassen	Skalenvarianz, wenn I-tem weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
inter_Item6	14.69	17.166	.498	.854
inter_Item7	15.03	18.574	.656	.811
inter_Item8	15.11	17.465	.711	.799
inter_Item9	14.74	18.154	.650	.811
inter_Item10	14.99	18.259	.646	.812
inter_Item11	15.13	18.420	.641	.813

Tabelle 210: Reliabilitätsanalyse – Fachinteresse gesamt

Reliabilitätsstatistiken				
Cronbachs Alpha		Anzahl der Items		
.897		11		
Item-Skala-Statistiken				
	Skalenmittelwert, wenn I-tem weggelassen	Skalenvarianz, wenn I-tem weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
inter_Item6	29.78	63.982	.501	.899
inter_Item7	30.12	66.260	.641	.888
inter_Item8	30.21	64.861	.654	.887
inter_Item9	29.83	66.350	.584	.891
inter_Item10	30.08	66.058	.610	.889
inter_Item11	30.22	65.659	.649	.887
inter_Item1	30.41	64.400	.727	.883
inter_Item2	29.11	64.681	.591	.891
inter_Item3	29.54	63.473	.638	.888
inter_Item4	30.35	63.680	.664	.886
inter_Item5	30.62	62.870	.729	.882

Tabelle 211: Fachinteresse – Test auf Normalverteilung

Tests auf Normalverteilung									
	Statistik	Kolmogorov-Smirnov ^a		Signifikanz	Statistik	Shapiro-Wilk		Signifikanz	
Fachinteresse	.073	df	107	Signifikanz	.200	.977	df	107	Signifikanz
a. Signifikanzkorrektur nach Lilliefors									

Tabelle 212: Fachinteresse – Test auf Unterscheidung der Gruppen – Parametrisch

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifi- kanz	T	df	Sig. (2- seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
Fachinteresse	Varianzen sind gleich	4.643	.033	1.285	105	.201	.22615	.17594	-.12271	.57501
	Varianzen sind nicht gleich			1.254	87.333	.213	.22615	.18028	-.13216	.58446

10.3.11. Statistiken zum Kapitel 7.1

Tabelle 213: Überprüfung der Unterscheidung von Pilot- und Hauptstudie – Reflexionskompetenztest

Tests der Zwischensubjekteffekte						
Abhängige Variable: Average Percentage Refl Post						
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	
Korrigiertes Modell	1033,093 ^a	2	516,547	1,252	,294	
Konstanter Term	17983,975	1	17983,975	43,587	,000	
PerReflPre	935,705	1	935,705	2,268	,138	
Studienzeitpnt_2	174,722	1	174,722	,423	,518	
Fehler	22692,960	55	412,599			
Gesamt	243055,556	58				
Korrigierte Gesamtvariation	23726,054	57				
a. R-Quadrat = ,044 (korrigiertes R-Quadrat = ,009)						

Tabelle 214: Überprüfung der Unterscheidung von Pilot- und Hauptstudie – Fachwissen

Tests der Zwischensubjekteffekte						
Abhängige Variable: Average Percentage Scores CK_Post						
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	
Korrigiertes Modell	3873,564 ^a	2	1936,782	6,273	,004	
Konstanter Term	6097,349	1	6097,349	19,749	,000	
PerCKPre	3854,393	1	3854,393	12,484	,001	
Studienzeitpnt_2	45,828	1	45,828	,148	,702	
Fehler	16980,864	55	308,743			
Gesamt	213434,903	58				
Korrigierte Gesamtvariation	20854,427	57				
a. R-Quadrat = ,186 (korrigiertes R-Quadrat = ,156)						

Tabelle 215: Überprüfung der Unterscheidung von Pilot- und Hauptstudie – Fachdidaktisches Wissen

Tests der Zwischensubjekteffekte					
Abhängige Variable: yScores PCK Experimentieren					
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.
Korrigiertes Modell	5412,281 ^a	2	2706,141	15,694	,000
Konstanter Term	877,759	1	877,759	5,090	,028
XScores	5054,906	1	5054,906	29,315	,000
Studienzeitpnt_2	222,844	1	222,844	1,292	,261
Fehler	9483,721	55	172,431		
Gesamt	308081,717	58			
Korrigierte Gesamtvariation	14896,002	57			

a. R-Quadrat = ,363 (korrigiertes R-Quadrat = ,340)

Tabelle 216: Überprüfung der Unterscheidung von Pilot- und Hauptstudie – Fachspezifische Überzeugungen

Tests der Zwischensubjekteffekte					
Abhängige Variable: yFachspezifischeÜberzeugungenGesamt					
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.
Korrigiertes Modell	2,158 ^a	2	1,079	28,819	,000
Konstanter Term	,110	1	,110	2,941	,092
yFachspezifischeÜberzeugungen	2,140	1	2,140	57,143	,000
Studienzeitpnt_2	,027	1	,027	,708	,404
Fehler	2,059	55	,037		
Gesamt	263,650	58			
Korrigierte Gesamtvariation	4,217	57			

a. R-Quadrat = ,512 (korrigiertes R-Quadrat = ,494)

Tabelle 217: Überprüfung der Unterscheidung von Pilot- und Hauptstudie – Selbstwirksamkeitserwartung

Tests der Zwischensubjekteffekte					
Abhängige Variable: ySelbstwirksamkeitserwartungGesamt					
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.
Korrigiertes Modell	4,351 ^a	2	2,175	31,865	,000
Konstanter Term	,459	1	,459	6,729	,012
ySelbstwirksamkeitserwartung	3,751	1	3,751	54,942	,000
Studienzeitpnt_2	,066	1	,066	,967	,330
Fehler	3,755	55	,068		
Gesamt	362,400	58			
Korrigierte Gesamtvariation	8,105	57			

a. R-Quadrat = ,537 (korrigiertes R-Quadrat = ,520)

10.3.12. Statistiken zum Kapitel 7.1.1

Tabelle 218: Reliabilitätsanalyse – Reflexionskompetenztest – Hauptstudie – 17 Items

Reliabilitätsstatistiken				
Cronbachs Alpha	Anzahl der Items			
.604	17			

Item-Skala-Statistiken				
	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
yRef_ Item1 Punkte	9.9821	7.109	.065	.611
yRef_ Item2 Punkte	10.1071	6.206	.417	.555
yRef_ Item3 Punkte	10.2679	6.200	.380	.560
yRef_ Item4 Punkte	10.6786	6.949	.219	.591
yRef_ Item5 Punkte	10.2143	6.790	.142	.604
yRef_ Item6 Punkte	9.8571	7.143	.142	.599
yRef_ Item7 Punkte	9.9464	7.106	.083	.607
yRef_ Item8 Punkte	9.9643	6.726	.266	.583
yRef_ Item9 Punkte	10.0536	6.633	.250	.585
yRef_ Item10 Punkte	10.0357	6.544	.301	.577
yRef_ Item11 Punkte	10.1607	6.356	.330	.570
yRef_ Item13 Punkte	9.8214	7.095	.279	.591
yRef_ Item14 Punkte	10.4286	7.195	-.008	.627
yRef_ Item16 Punkte	10.4107	7.046	.047	.619
yRef_ Item19 Punkte	10.1786	6.222	.384	.560
yRef_ Item20 Punkte	10.3393	6.519	.249	.585
yRef_ Item21 Punkte	10.1250	6.475	.290	.578

Tabelle 219: Reliabilitätsanalyse – Reflexionskompetenztest – Hauptstudie – 12 Items

Reliabilitätsstatistiken	
Cronbachs Alpha	Anzahl der Items

Item-Skala-Statistiken				
	Skalenmittelwert, wenn I-tem weggelassen	Skalenvarianz, wenn I-tem weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
yRefI_Item2 Punkte	6.8036	4.815	.413	.612
yRefI_Item3 Punkte	6.9643	4.908	.326	.628
yRefI_Item4 Punkte	7.3750	5.475	.218	.646
yRefI_Item5 Punkte	6.9107	5.210	.189	.655
yRefI_Item8 Punkte	6.6607	5.246	.279	.637
yRefI_Item9 Punkte	6.7500	5.100	.293	.634
yRefI_Item10 Punkte	6.7321	5.072	.319	.630
yRefI_Item11 Punkte	6.8571	4.779	.409	.612
yRefI_Item13 Punkte	6.5179	5.672	.205	.649
yRefI_Item19 Punkte	6.8750	4.766	.410	.611
yRefI_Item20 Punkte	7.0357	5.162	.209	.651
yRefI_Item21 Punkte	6.8214	5.058	.281	.637

Tabelle 220: Reliabilitätsanalyse – Reflexionskompetenztest – Selbstreflexions- und Kontrollgruppe

Reliabilitätsstatistiken	
Cronbachs Alpha	Anzahl der Items
.726	15

Item-Skala-Statistiken				
	Skalenmittelwert, wenn I-tem weggelassen	Skalenvarianz, wenn I-tem weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
yRefI_Item1 Punkte	8.9167	8.193	.172	.726
yRefI_Item2 Punkte	9.1111	7.244	.467	.694
yRefI_Item3 Punkte	9.2778	7.006	.539	.684
yRefI_Item4 Punkte	9.6667	8.229	.249	.719
yRefI_Item5 Punkte	9.1667	7.857	.214	.726
yRefI_Item6 Punkte	8.8333	8.371	.159	.726
yRefI_Item7 Punkte	8.9444	8.054	.216	.723
yRefI_Item8 Punkte	8.9444	8.054	.216	.723
yRefI_Item9 Punkte	9.0000	7.714	.328	.712
yRefI_Item10 Punkte	9.0278	7.513	.396	.704
yRefI_Item11 Punkte	9.1667	7.229	.457	.695
yRefI_Item13 Punkte	8.8056	8.218	.327	.716
yRefI_Item19 Punkte	9.1944	7.304	.422	.700
yRefI_Item20 Punkte	9.3333	7.429	.377	.706
yRefI_Item21 Punkte	9.1111	7.587	.329	.712

Tabelle 221: Reflexionskompetenztest – Test auf Normalverteilung

Tests auf Normalverteilung						
Statistik	Kolmogorov-Smirnov ^a		Signifikanz	Statistik	Shapiro-Wilk	
	df				df	Signifikanz
Average Percentage Refl Pre	.181	58	.000	.914	58	.001
Average Percentage Refl Post	.145	58	.004	.950	58	.017
Average Percentage Refl FollowUp	.096	58	.200 [*]	.975	58	.268

*. Dies ist eine untere Grenze der echten Signifikanz.

a. Signifikanzkorrektur nach Lilliefors

Tabelle 222: Reflexionskompetenztest – Mittelwertsvergleich – Nichtparametrisch

Statistik für Test ^a			
	Average Percentage Refl Post - Average Percentage Refl Pre	Average Percentage Refl FollowUp - Average Percentage Refl Pre	
Z		-6.409 ^b	-6.014 ^b
Asymptotische Signifikanz (2-seitig)		.000	.000

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 223: Reflexionskompetenztest – Mittelwertsvergleich – Nichtparametrisch – Feedbackgruppe

Statistik für Test ^a			
	Average Percentage Refl Post - Average Percentage Refl Pre	Average Percentage Refl FollowUp - Average Percentage Refl Pre	
Z		-3.936 ^b	-3.734 ^b
Asymptotische Signifikanz (2-seitig)		.000	.000

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 224: Reflexionskompetenztest – Mittelwertsvergleich – Nichtparametrisch – Selbstreflexionsgruppe

Statistik für Test^a

10. Anhang

	Average Percentage Refl Post - Average Percentage Refl Pre	Average Percentage Refl FollowUp - Average Percentage Refl Pre
Z	-3.734 ^b	-3.316 ^b
Asymptotische Signifikanz (2-seitig)	.000	.001
a. Wilcoxon-Test		
b. Basiert auf negativen Rängen.		

Tabelle 225: Reflexionskompetenztest – Mittelwertsvergleich – Nichtparametrisch – Kontrollgruppe

Statistik für Test ^a		
	Average Percentage Refl Post - Average Percentage Refl Pre	Average Percentage Refl FollowUp - Average Percentage Refl Pre
Z	-3.513 ^b	-3.275 ^b
Asymptotische Signifikanz (2-seitig)	.000	.001
a. Wilcoxon-Test		
b. Basiert auf negativen Rängen.		

Tabelle 226: Reflexionskompetenztest – Mittelwertsvergleich – Parametrisch

Test bei gepaarten Stichproben									
Gepaarte Differenzen									
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	Refl Pre - Refl Post	-36.35057	20.74973	2.72457	-41.80644	-30.89471	-13.342	57	.000
Paaren 2	Refl Pre - Refl FollowUp	-24.85632	20.44172	2.68413	-30.23120	-19.48145	-9.260	57	.000

Tabelle 227: Reflexionskompetenztest – Mittelwertsvergleich – Parametrisch – Feedbackgruppe

Test bei gepaarten Stichproben									
Gepaarte Differenzen									
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	Refl Pre – Refl Post	-43.33333	17.01393	3.80443	-51.29610	-35.37057	-11.390	19	.000
Paaren 2	Refl Pre – Refl FollowUp	-29.58333	18.43175	4.12147	-38.20966	-20.95701	-7.178	19	.000

Tabelle 228: Reflexionskompetenztest – Mittelwertsvergleich – Parametrisch – Selbstreflexionsgruppe

Test bei gepaarten Stichproben									
Gepaarte Differenzen									
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	Refl Pre - Refl Post	-27.38095	19.02901	4.15247	-36.04285	-18.71905	-6.594	20	.000
Paaren 2	Refl Pre - Refl FollowUp	-17.46032	16.22429	3.54043	-24.84552	-10.07511	-4.932	20	.000

Tabelle 229: Reflexionskompetenztest – Mittelwertsvergleich – Parametrisch – Kontrollgruppe

Test bei gepaarten Stichproben									
Gepaarte Differenzen									
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	Refl Pre - Refl Post	-39.21569	23.71062	5.75067	-51.40656	-27.02481	-6.819	16	.000
Paaren 2	Refl Pre - Refl FollowUp	-28.43137	25.35492	6.14947	-41.46767	-15.39508	-4.623	16	.000

Tabelle 230: Ergebnisse des Reflexionskompetenztests – Gruppenvergleiche parametrisch – Gesamtstudie

Gruppe	Messzeitpunkte	t	p	d
Feedbackgruppe	Pre-Post	$t(19) = 11.39$	< .001	3.6
	Pre-FU	$t(19) = 7.18$	< .001	2.1
Selbstreflexionsgruppe	Pre-Post	$t(20) = 6.59$	< .001	1.6
	Pre-FU	$t(20) = 4.93$	< .001	1.2

Kontrollgruppe	Pre-Post	$t(16) = 6.82$	$< .001$	2.2
	Pre-FU	$t(16) = 4.62$	$< .001$	1.7

10.3.13. Statistiken zum Kapitel 7.1.2

Gesamtpunktzahl

Tabelle 231: Kodiermanual – Gesamtpunktzahl – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest		xGesamtpunktzahl	yGesamtpunktzahl	zGesamtpunktzahl
N		58	58	58
Parameter der Normalverteilung ^{a,b}	M	63.1379	63.4310	76.1379
	Std.-Abweichung	29.72414	30.50862	36.31995
Extremste Differenzen	Absolut	.102	.122	.179
	Positiv	.102	.122	.179
	Negativ	-.089	-.106	-.078
Statistik für Test		.102	.122	.179
Asymptotische Signifikanz (2-seitig)		.200 ^{c,d}	.031 ^c	.000 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

d. Dies ist eine untere Grenze der echten Signifikanz.

Gesamtpunktzahl - Wissenszuwachs

Tabelle 232: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Nichtparametrisch – Gesamtstichprobe

Statistik für Test ^a		yGesamtpunktzahl - xGesamtpunktzahl	zGesamtpunktzahl - xGesamtpunktzahl
Z		-.200 ^b	-3.660 ^b
Asymptotische Signifikanz (2-seitig)		.842	.000

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 233: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Parametrisch – Gesamtstichprobe

Test bei gepaarten Stichproben									
				Gepaarte Differenzen					
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	xGesamtpunktzahl - yGesamtpunktzahl	-.29310	19.63081	2.57765	-5.45476	4.86856	-.114	57	.910
Paaren 2	xGesamtpunktzahl - zGesamtpunktzahl	-13.00000	25.49234	3.34731	-19.70287	-6.29713	-3.884	57	.000

Tabelle 234: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Nichtparametrisch – Feedbackgruppe

Statistik für Test ^a		yGesamtpunktzahl - xGesamtpunktzahl	zGesamtpunktzahl - xGesamtpunktzahl
Z		-1.550 ^b	-2.241 ^c
Asymptotische Signifikanz (2-seitig)		.121	.025

a. Wilcoxon-Test

b. Basiert auf positiven Rängen.

c. Basiert auf negativen Rängen.

Tabelle 235: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Nichtparametrisch – Selbstreflexionsgruppe

Statistik für Test ^a		yGesamtpunktzahl - xGesamtpunktzahl	zGesamtpunktzahl - xGesamtpunktzahl
Z		-1.137 ^b	-1.423 ^b
Asymptotische Signifikanz (2-seitig)		.255	.155

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 236: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Nichtparametrisch – Kontrollgruppe

Statistik für Test ^a		
	yGesamtpunktzahl - xGesamtpunktzahl	zGesamtpunktzahl - xGesamtpunktzahl
Z	-1.137 ^b	-1.423 ^b
Asymptotische Signifikanz (2-seitig)	.255	.155

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 237: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Parametrisch – Feedbackgruppe

Test bei gepaarten Stichproben								
		Gepaarte Differenzen						
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz			
					Untere	Obere	T	Sig. (2-seitig)
Paaren 1	xGesamtpunktzahl - yGesamtpunktzahl	6.70000	19.37741	4.33292	-2.36891	15.76891	1.546	.139
Paaren 2	xGesamtpunktzahl - zGesamtpunktzahl	-13.70000	26.59858	5.94762	-26.14852	-1.25148	-2.303	.033

Tabelle 238: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Parametrisch – Selbstreflexionsgruppe

Test bei gepaarten Stichproben								
		Gepaarte Differenzen						
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz			
					Untere	Obere	T	Sig. (2-seitig)
Paaren 1	xGesamtpunktzahl - yGesamtpunktzahl	-3.57143	20.54890	4.48414	-12.92517	5.78232	-.796	.435
Paaren 2	xGesamtpunktzahl - zGesamtpunktzahl	-14.80952	26.49834	5.78241	-26.87142	-2.74763	-2.561	.019

Tabelle 239: Kodiermanual – Gesamtpunktzahl – Mittelwertsvergleich – Parametrisch – Kontrollgruppe

Test bei gepaarten Stichproben								
		Gepaarte Differenzen						
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz			
					Untere	Obere	T	Sig. (2-seitig)
Paaren 1	xGesamt - yGesamt	-4.47059	17.47540	4.23841	-13.45561	4.51443	-1.055	.307
Paaren 2	xGesamt - zGesamt	-9.94118	24.13367	5.85327	-22.34956	2.46721	-1.698	.109

Reflexionsbreite

Tabelle 240: Kodiermanual – Reflexionsbreite – Test auf Normalverteilung

Tests auf Normalverteilung						
Statistik	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	df	Signifikanz		df	Signifikanz	
xReflexionsbreite	.147	58	.003	.952	58	.023
yReflexionsbreite	.179	58	.000	.948	58	.014
zReflexionsbreite	.137	58	.009	.909	58	.000

a. Signifikanzkorrektur nach Lilliefors

Reflexionsbreite - Wissenszuwachs:

Tabelle 241: Kodiermanual – Reflexionsbreite – Mittelwertsvergleich – Nichtparametrisch – Gesamtstichprobe

Statistik für Test ^a			
	yReflexionsbreite - xReflexionsbreite	zReflexionsbreite - xReflexionsbreite	zReflexionsbreite - yReflexionsbreite
Z	-.729 ^b	-1.474 ^b	-1.020 ^b
Asymptotische Signifikanz (2-seitig)	.466	.140	.308

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 242: Kodiermanual – Reflexionsbreite – Mittelwertsvergleich – Nichtparametrisch – Feedbackgruppe

	Statistik für Test ^a		
	yReflexionsbreite - xReflexionsbreite	zReflexionsbreite - xReflexionsbreite	zReflexionsbreite - yReflexionsbreite
Z	-.100 ^b	-1.615 ^c	-1.574 ^c
Asymptotische Signifikanz (2-seitig)	.921	.106	.116

a. Wilcoxon-Test

b. Basiert auf positiven Rängen.

c. Basiert auf negativen Rängen.

Tabelle 243: Kodiermanual – Reflexionsbreite – Mittelwertsvergleich – Nichtparametrisch – Selbstreflexionsgruppe

	Statistik für Test ^a		
	yReflexionsbreite - xReflexionsbreite	zReflexionsbreite - xReflexionsbreite	zReflexionsbreite - yReflexionsbreite
Z	-.440 ^b	-1.081 ^b	-.944 ^b
Asymptotische Signifikanz (2-seitig)	.660	.280	.345

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 244: Kodiermanual – Reflexionsbreite – Mittelwertsvergleich – Nichtparametrisch – Kontrollgruppe

	Statistik für Test ^a		
	yReflexionsbreite - xReflexionsbreite	zReflexionsbreite - xReflexionsbreite	zReflexionsbreite - yReflexionsbreite
Z	-.750 ^b	-.026 ^c	-.966 ^c
Asymptotische Signifikanz (2-seitig)	.453	.979	.334

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

c. Basiert auf positiven Rängen.

Reflexionstiefe:

Tabelle 245: Verteilung der Kategorien der Reflexionstiefe

	Deskriptive Statistik				
	N	Min.	Max.	M	Std.-Abweichung
x1T_Beschreibung	51	1	16	5.27	3.945
x1T_Rechtfertigung	26	1	6	2.19	1.600
x1T_Abwägen	10	1	7	2.00	1.826
x1T_Ziele	25	1	10	2.00	1.893
x2T_Bewertung_negativ	57	1	46	8.46	7.177
x2T_Bewertung_positiv	53	1	17	6.08	4.385
x2T_Rechtfertigung	56	1	36	7.16	6.225
x2T_Abwägen	40	1	14	2.80	2.483
x2T_Ziele	28	1	8	2.46	1.835
x3T_Alternative	36	1	14	2.00	2.255
x3T_Rechtfertigung	16	1	4	1.50	.816
x3T_Abwägen	11	1	10	2.00	2.683
x3T_Ziele	0				
x4T_Konsequenzen_Verbesserungsvorschläge	58	2	30	11.43	5.120
x4T_Rechtfertigungen	56	1	35	8.66	5.609
x4T_Abwägen	33	1	10	2.00	1.750
x4T_Ziele	14	1	2	1.29	.469
x5_1Schülerperspektive	49	1	13	3.04	2.406
x5_2Theorieperspektive	49	1	17	5.02	3.733
x5_3Fachperspektive	8	1	2	1.13	.354
x5_4PolitischSozialePerspektive	3	1	4	2.00	1.732
x6Literaturbezug	9	1	4	1.56	1.014
y1T_Beschreibung	52	1	25	5.48	4.621
y1T_Rechtfertigung	28	1	17	2.71	3.196
y1T_Abwägen	7	1	4	1.57	1.134
y1T_Ziele	22	1	8	2.23	1.926
y2T_Bewertung_negativ	56	1	30	5.73	4.856
y2T_Bewertung_positiv	57	1	25	9.23	6.173
y2T_Rechtfertigung	55	1	35	9.89	8.004
y2T_Abwägen	44	1	19	3.05	3.065
y2T_Ziele	34	1	5	1.85	1.105
y3T_Alternative	30	1	14	2.10	2.454
y3T_Rechtfertigung	15	1	4	1.73	1.100
y3T_Abwägen	9	1	12	2.67	3.571
y3T_Ziele	1	1	1	1.00	.
y4T_Konsequenz	57	2	24	9.39	4.378
y4T_Rechtfertigung	57	1	22	7.14	4.886
y4T_Abwägen	32	1	14	2.00	2.410
y4T_Ziele	9	1	3	1.44	.726
y5_1Schülerperspektive	50	1	10	3.00	2.416
y5_2Theorieperspektive	54	1	12	4.50	3.076
y5_3Fachperspektive	12	1	2	1.08	.289

10. Anhang

y5_4PolitischSozialePerspektive	6	1	1	1.00	.000
y6Literaturbezug	13	1	6	2.15	1.463
z1T_Beschreibung	55	1	20	6.65	4.850
z1T_Rechtfertigung	30	1	23	3.27	4.042
z1T_Abwägen	8	1	7	3.13	2.100
z1T_Ziele	20	1	7	2.25	1.773
z2T_Bewertung_negativ	52	1	22	5.23	4.250
z2T_Bewertung_positiv	57	1	41	13.60	9.140
z2T_Rechtfertigung	56	1	50	14.45	10.579
z2T_Abwägen	48	1	14	3.83	3.191
z2T_Ziele	28	1	8	2.79	2.114
z3T_Alternative	27	1	7	2.52	1.762
z3T_Rechtfertigung	17	1	5	2.41	1.417
z3T_Abwägen	14	1	5	2.43	1.222
z3T_Ziele	0				
z4T_Konsequenz	55	1	18	7.35	4.270
z4T_Rechtfertigung	54	1	14	6.00	3.938
z4T_Abwägen	32	1	8	2.09	1.692
z4T_Ziele	8	1	5	1.50	1.414
z5_1Schülerperspektive	50	1	12	3.80	2.688
z5_2Theorieperspektive	56	1	16	7.25	3.502
z5_3Fachperspektive	15	1	3	1.60	.737
z5_4PolitischSozialePerspektive	6	1	1	1.00	.000
z6Literaturbezug	43	1	12	3.70	2.891
Gültige Werte (Listenweise)	0				

Tabelle 246: Kodiermanual – Reflexionstiefe – Test auf Normalverteilung

	Tests auf Normalverteilung					
	Statistik	Kolmogorov-Smirnov ^a	Signifikanz	Statistik	Shapiro-Wilk	Signifikanz
xReflexionstiefe	.108	58	.088	.888	58	.000
yReflexionstiefe	.130	58	.016	.911	58	.000
zReflexionstiefe	.166	58	.000	.922	58	.001

a. Signifikanzkorrektur nach Lilliefors

Tabelle 247: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Nichtparametrisch – Gesamtstichprobe

	Statistik für Test ^a		
	yReflexionstiefe - xReflexionstiefe	zReflexionstiefe - xReflexionstiefe	zReflexionstiefe - yReflexionstiefe
Z	-.180 ^b	-3.740 ^b	-3.423 ^b
Asymptotische Signifikanz (2-seitig)	.858	.000	.001

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 248: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Parametrisch – Gesamtstichprobe

Test bei gepaarten Stichproben										
		Gepaarte Differenzen								
		Mittelwert	Std.-Abwei- chung	Standardfeh- ler des Mittel- wertes	95% Konfidenzintervall der Dif- ferenz		T	df	Sig. (2-seitig)	
Paaren 1	xReflexionstiefe - yReflexi- onstiefe	-1.17241	19.14042	2.51326	-5.20513	4.86030	-0.069	57	.946	
Paaren 2	xReflexionstiefe - zReflexi- onstiefe	-12.65517	24.72980	3.24718	-19.15754	-6.15280	-3.897	57	.000	
Paaren 3	yReflexionstiefe - zReflexi- onstiefe	-12.48276	26.63940	3.49792	-19.48723	-5.47829	-3.569	57	.001	

Tabelle 249: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Nichtparametrisch – Feedbackgruppe

	Statistik für Test ^a		
	yReflexionstiefe - xReflexionstiefe	zReflexionstiefe - xReflexionstiefe	zReflexionstiefe - yReflexionstiefe
Z	-1.438 ^b	-2.279 ^c	-3.249 ^c
Asymptotische Signifikanz (2-seitig)	.150	.023	.001

a. Wilcoxon-Test

b. Basiert auf positiven Rängen.

c. Basiert auf negativen Rängen.

Tabelle 250: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Parametrisch – Feedbackgruppe

Test bei gepaarten Stichproben

		Gepaarte Differenzen						Sig. (2-seitig)	
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	
					Untere	Obere			
Paaren 1	xReflexionstiefe - yReflexionstiefe	6.70000	18.78997	4.20157	-2.09398	15.49398	1.595	19	.127
Paaren 2	xReflexionstiefe - zReflexionstiefe	-13.10000	26.22554	5.86421	-25.37393	-.82607	-2.234	19	.038
Paaren 3	yReflexionstiefe - zReflexionstiefe	-19.80000	25.81024	5.77134	-31.87956	-7.72044	-3.431	19	.003

Tabelle 251: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Nichtparametrisch – Selbstreflexionsgruppe

Statistik für Test ^a			
	yReflexionstiefe - xReflexionstiefe	zReflexionstiefe - xReflexionstiefe	zReflexionstiefe - yReflexionstiefe
Z	-.730 ^b	-2.468 ^b	-1.965 ^b
Asymptotische Signifikanz (2-seitig)	.465	.014	.049

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 252: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Parametrisch – Selbstreflexionsgruppe

Test bei gepaarten Stichproben									
		Gepaarte Differenzen							
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	xReflexionstiefe - yReflexionstiefe	-3.47619	20.31408	4.43290	-12.72305	5.77067	-.784	20	.442
Paaren 2	xReflexionstiefe - zReflexionstiefe	-14.38095	25.42140	5.54740	-25.95264	-2.80927	-2.592	20	.017
Paaren 3	yReflexionstiefe - zReflexionstiefe	-10.90476	28.33003	6.18212	-23.80044	1.99091	-1.764	20	.093

Tabelle 253: Kodiermanual - Reflexionstiefe - Mittelwertsvergleich - Nichtparametrisch - Kontrollgruppe

Statistik für Test ^a			
	yReflexionstiefe - xReflexionstiefe	zReflexionstiefe - xReflexionstiefe	zReflexionstiefe - yReflexionstiefe
Z	-1.090 ^b	-1.492 ^b	-.900 ^b
Asymptotische Signifikanz (2-seitig)	.276	.136	.368

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 254: Kodiermanual – Reflexionstiefe – Mittelwertsvergleich – Parametrisch – Kontrollgruppe

Test bei gepaarten Stichproben									
		Gepaarte Differenzen							
		Mittelwert	Std.-Abwei- chung	Standardfeh- ler des Mittel- wertes	95% Konfidenzintervall der Dif- ferenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	xReflexionstiefe - yRefle- xionstiefe	-4.17647	16.72287	4.05589	-12.77458	4.42164	-1.030	16	.318
Paaren 2	xReflexionstiefe - zRefle- xionstiefe	-10.00000	23.29163	5.64905	-21.97545	1.97545	-1.770	16	.096
Paaren 3	yReflexionstiefe - zRefle- xionstiefe	-5.82353	24.83001	6.02216	-18.58994	6.94288	-.967	16	.348

10.3.14. Statistiken zum Kapitel 7.2.1

PCK-Test

Tabelle 255: PCK-Test – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest				
		xScores PCK Experimentieren	yScores PCK Experimentieren	zScores PCK Experimentieren
N		58	58	58
Parameter der Normalverteilung ^{a,b}	Mittelwert	67.8766	71.0980	69.3285
	Std.-Abweichung	12.96496	16.16581	11.98920
Extremste Differenzen	Absolut	.137	.144	.091

10. Anhang

	Positiv	.087	.076	.056
	Negativ	-.137	-.144	-.091
Statistik für Test		.137	.144	.091
Asymptotische Signifikanz (2-seitig)		.008 ^c	.004 ^c	.200 ^{c,d}

- a. Die zu testende Verteilung ist eine Normalverteilung.
b. Aus den Daten berechnet.
c. Signifikanzkorrektur nach Lilliefors.
d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 256: PCK-Test – Mittelwertsvergleich – Nichtparametrisch – Gesamtstichprobe

Statistik für Test ^a		
	yScores PCK Experimentieren - xScores PCK Experimentieren	zScores PCK Experimentieren - xScores PCK Experimentieren
Z	-1.817 ^b	-.833 ^b
Asymptotische Signifikanz (2-seitig)	.069	.405

- a. Wilcoxon-Test
b. Basiert auf negativen Rängen.

Tabelle 257: PCK-Test – Mittelwertsvergleich – Parametrisch – Gesamtstichprobe

Test bei gepaarten Stichproben									
		Gepaarte Differenzen			95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	xScores PCK Experimentieren - yScores PCK Experimentieren	-3.22142	13.49110	1.77147	-6.76872	.32589	-1.819	57	.074
Paaren 2	xScores PCK Experimentieren - zScores PCK Experimentieren	-1.45191	11.16672	1.46626	-4.38804	1.48423	-.990	57	.326

CK-Test

Tabelle 258: CK-Test – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest				
		Average Percentage Scores CK_Pre	Average Percentage Scores CK_Post	Average Percentage Scores CK_FollowUp
N		58	58	58
Parameter der Normalverteilung ^{a,b}	Mittelwert	39.9274	57.6225	47.1869
	Std.-Abweichung	13.80353	19.12765	15.39885
Extremste Differenzen	Absolut	.132	.110	.094
	Positiv	.107	.101	.094
	Negativ	-.132	-.110	-.067
Statistik für Test		.132	.110	.094
Asymptotische Signifikanz (2-seitig)		.014 ^c	.079 ^c	.200 ^{c,d}

- a. Die zu testende Verteilung ist eine Normalverteilung.
b. Aus den Daten berechnet.
c. Signifikanzkorrektur nach Lilliefors.
d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 259: CK-Test – Mittelwertsvergleich – Nichtparametrisch – Gesamtstichprobe

Statistik für Test ^a		
	Average Percentage Scores CK_Post - Average Percentage Scores CK_Pre	Average Percentage Scores CK_FollowUp - Average Percentage Scores CK_Pre
Z	-5.536 ^b	-3.452 ^b
Asymptotische Signifikanz (2-seitig)	.000	.001

- a. Wilcoxon-Test
b. Basiert auf negativen Rängen.

Tabelle 260: CK-Test – Mittelwertsvergleich – Parametrisch – Gesamtstichprobe

Test bei gepaarten Stichproben									
		Gepaarte Differenzen							
			Std.-Abwei- chung	Standardfeh- ler des Mit- telwertes	95% Konfidenzintervall der Dif- ferenz		T	df	Sig. (2-seitig)
		Mittelwert			Untere	Obere			
Paaren 1	Average Percentage Scores CK_Pre - Average Percentage Scores CK_Post	-17.69510	18.17066	2.38592	-22.47283	-12.91737	-7.416	57	.000
Paaren 2	Average Percentage Scores CK_Pre - Average Percentage Scores CK_FollowUp	-7.25953	15.13984	1.98796	-11.24035	-3.27871	-3.652	57	.001

Korrelationen der residualen Lernzuwächse

Gesamtstichprobe

Tabelle 261: Residualer Lernzuwachs – Test auf Normalverteilung

		Kolmogorov-Smirnov-Anpassungstest							
		SR PCK Gesamt	SR CK Gesamt	SR ReflBreite Gesamt	SR ReflTiefe Gesamt	SR ReflPerspektive Gesamt	SR ReflLiteratur Gesamt	SR ReflGesamt Gesamt	SR ReflTest Gesamt
N		58	58	58	58	57	57	58	58
Parameter der Normalverteilung ^{a,b}	Mittelwert	.0000000	.0000000	.0000000	.0000000	.0000000	.0000000	.0000000	.0000000
	Std.-Abweichung	.99118926	.99118926	.99118926	.99118926	.99103121	.99103121	.99118926	.99118926
Extremste Differenzen	Absolut	.128	.129	.090	.162	.108	.131	.151	.100
	Positiv	.089	.129	.061	.162	.108	.131	.151	.079
	Negativ	-.128	-.091	-.090	-.082	-.049	-.064	-.064	-.100
Statistik für Test		.128	.129	.090	.162	.108	.131	.151	.100
Asymptotische Signifikanz (2-seitig)		.019 ^c	.018 ^c	.200 ^{c,d}	.001 ^c	.095 ^c	.017 ^c	.002 ^c	.200 ^{c,d}

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 262: Parametrische Korrelationen

		Korrelationen							
		SR PCK Gesamt	SR CK Gesamt	SR ReflBreite Gesamt	SR ReflTiefe Gesamt	SR ReflPerspektive Gesamt	SR ReflLiteratur Gesamt	SR ReflGesamt Gesamt	SR ReflTest Gesamt
SR PCK Gesamt	Korrelation nach Pearson	1	.312 [*]	-.078	.043	.002	-.038	.038	.178
	Signifikanz (2-seitig)		.017	.560	.749	.986	.778	.779	.182
	N	58	58	58	58	57	57	58	58
SR CK Gesamt	Korrelation nach Pearson	.312 [*]	1	-.026	.033	.094	.066	.028	.182
	Signifikanz (2-seitig)	.017		.848	.806	.486	.624	.833	.171
	N	58	58	58	58	57	57	58	58
SR ReflBreite Gesamt	Korrelation nach Pearson	-.078	-.026	1	.430 ^{**}	.474 ^{**}	.442 ^{**}	.479 ^{**}	.063
	Signifikanz (2-seitig)	.560	.848		.001	.000	.001	.000	.639
	N	58	58	58	58	57	57	58	58
SR ReflTiefe Gesamt	Korrelation nach Pearson	.043	.033	.430 ^{**}	1	.674 ^{**}	.727 ^{**}	.998 ^{**}	.183
	Signifikanz (2-seitig)	.749	.806	.001		.000	.000	.000	.169
	N	58	58	58	58	57	57	58	58
SR ReflPerspektive Gesamt	Korrelation nach Pearson	.002	.094	.474 ^{**}	.674 ^{**}	1	.915 ^{**}	.676 ^{**}	.154
	Signifikanz (2-seitig)	.986	.486	.000	.000		.000	.000	.252
	N	57	57	57	57	57	57	57	57
SR ReflLiteratur Gesamt	Korrelation nach Pearson	-.038	.066	.442 ^{**}	.727 ^{**}	.915 ^{**}	1	.725 ^{**}	.179
	Signifikanz (2-seitig)	.778	.624	.001	.000	.000		.000	.182
	N	57	57	57	57	57	57	57	57
SR ReflGesamt Gesamt	Korrelation nach Pearson	.038	.028	.479 ^{**}	.998 ^{**}	.676 ^{**}	.725 ^{**}	1	.183
	Signifikanz (2-seitig)	.779	.833	.000	.000	.000	.000		.170
	N	58	58	58	58	57	57	58	58
SR ReflTest Gesamt	Korrelation nach Pearson	.178	.182	.063	.183	.154	.179	.183	1
	Signifikanz (2-seitig)	.182	.171	.639	.169	.252	.182	.170	
	N	58	58	58	58	57	57	58	58

*. Die Korrelation ist auf dem Niveau von 0.05 (2-seitig) signifikant.

**. Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant.

Tabelle 263: Nichtparametrische Korrelationen

Korrelationen

10. Anhang

			SR PCK Gesamt	SR CK Gesamt	SR Re- flBreite Gesamt	SR Refl- Tiefe Ge- samt	SR Refl- Perspek- tive Ge- samt	SR ReflLi- teratur Gesamt	SR Refl- Gesamt Gesamt	SR Refl- Test Ge- samt
Spearman- Rho	SR PCK Gesamt	Korrelationsko- effizient	1.000	.302*	-.106	-.071	.060	-.029	-.053	.214
		Sig. (2-seitig)	.	.021	.427	.599	.655	.828	.693	.107
		N	58	58	58	58	57	57	58	58
	SR CK Gesamt	Korrelationsko- effizient	.302*	1.000	-.077	.030	.094	.035	.038	.186
		Sig. (2-seitig)	.021	.	.563	.820	.485	.796	.778	.163
		N	58	58	58	58	57	57	58	58
	SR ReflBreite Ge- samt	Korrelationsko- effizient	-.106	-.077	1.000	.472**	.459**	.436**	.533**	.057
		Sig. (2-seitig)	.427	.563	.	.000	.000	.001	.000	.673
		N	58	58	58	58	57	57	58	58
	SR ReflTiefe Ge- samt	Korrelationsko- effizient	-.071	.030	.472**	1.000	.664**	.688**	.993**	.118
		Sig. (2-seitig)	.599	.820	.000	.	.000	.000	.000	.380
		N	58	58	58	58	57	57	58	58
	SR ReflPerspek- tive Gesamt	Korrelationsko- effizient	.060	.094	.459**	.664**	1.000	.928**	.653**	.090
		Sig. (2-seitig)	.655	.485	.000	.000	.	.000	.000	.504
		N	57	57	57	57	57	57	57	57
	SR ReflLiteratur Gesamt	Korrelationsko- effizient	-.029	.035	.436**	.688**	.928**	1.000	.674**	.123
		Sig. (2-seitig)	.828	.796	.001	.000	.000	.	.000	.363
		N	57	57	57	57	57	57	57	57
	SR ReflGesamt Gesamt	Korrelationsko- effizient	-.053	.038	.533**	.993**	.653**	.674**	1.000	.120
		Sig. (2-seitig)	.693	.778	.000	.000	.000	.000	.	.369
		N	58	58	58	58	57	57	58	58
	SR ReflTest Ge- samt	Korrelationsko- effizient	.214	.186	.057	.118	.090	.123	.120	1.000
		Sig. (2-seitig)	.107	.163	.673	.380	.504	.363	.369	.
		N	58	58	58	58	57	57	58	58

*. Die Korrelation ist auf dem 0.05 Niveau signifikant (zweiseitig).

** Die Korrelation ist auf dem 0.01 Niveau signifikant (zweiseitig).

10.3.15. Statistiken zum Kapitel 7.2.2

Tabelle 264: Reliabilitäten – Zielorientierungen

Reliabilitätsstatistiken	
Cronbachs Alpha	Anzahl der Items
.665	16

Tabelle 265: Reliabilitäten – Fachspezifische Überzeugungen

Reliabilitätsstatistiken	
Cronbachs Alpha	Anzahl der Items
.529	10
.776	21
.499	8

Tabelle 266: Reliabilitäten – Selbstwirksamkeitserwartungen

Reliabilitätsstatistiken	
Cronbachs Alpha	Anzahl der Items
.665	5
.735	4
.767	8
.685	5

Tabelle 267: Zielorientierungen – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest			
		xBestrebungen im Stu- dium	zBestrebungen im Stu- dium
N		58	58
Parameter der Normalverteilung ^{a,b}	Mittelwert	1.9063	1.9386
	Std.-Abweichung	.34675	.33799
Extremste Differenzen	Absolut	.107	.083
	Positiv	.107	.083
	Negativ	-.092	-.082
Statistik für Test		.107	.083
Asymptotische Signifikanz (2-seitig)		.099 ^c	.200 ^{c,d}

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.
d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 268: Fachspezifische Überzeugungen – Test auf Normalverteilung

		Kolmogorov-Smirnov-Anpassungstest					
		xChemie als Wissenschaft	xUnterrichten von Chemie in der Schule	xLernen von Chemie in der Schule	zChemie als Wissenschaft	zUnterrichten von Chemie in der Schule	zLernen von Chemie in der Schule
N		58	58	58	58	58	58
Parameter der Normalverteilung ^{a,b}	Mittelwert	2.2552	2.1010	2.4871	2.1310	2.0328	2.3821
	Std.-Abweichung	.37142	.27523	.36099	.34651	.31577	.33538
Extremste Differenzen	Absolut	.100	.079	.105	.096	.124	.131
	Positiv	.073	.079	.105	.064	.122	.131
	Negativ	-.100	-.046	-.094	-.096	-.124	-.120
Statistik für Test		.100	.079	.105	.096	.124	.131
Asymptotische Signifikanz (2-seitig)		.200 ^{c,d}	.200 ^{c,d}	.178 ^c	.200 ^{c,d}	.028 ^c	.015 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 269: Selbstwirksamkeitserwartung – Test auf Normalverteilung

		Kolmogorov-Smirnov-Anpassungstest							
		xPlanung von Experimenten	xDurchführung von Experimenten	xdidaktische Rekonstruktion unterrichtsrelevanter Inhalte	xUmsetzung der didaktischen Rekonstruktion im Unterricht	zPlanung von Experimenten	zDurchführung von Experimenten	zdidaktische Rekonstruktion unterrichtsrelevanter Inhalte	zUmsetzung der didaktischen Rekonstruktion im Unterricht
N		58	58	58	58	58	58	58	58
Parameter der Normalverteilung ^{a,b}	Mittelwert	2.8862	2.7457	2.6078	2.5103	2.5655	2.4440	2.4009	2.2862
	Std.-Abweichung	.48862	.49225	.41520	.55209	.42738	.45171	.43032	.43223
Extremste Differenzen	Absolut	.169	.113	.087	.160	.116	.140	.094	.138
	Positiv	.169	.113	.087	.160	.116	.140	.094	.138
	Negativ	-.108	-.111	-.086	-.084	-.110	-.127	-.088	-.138
Statistik für Test		.169	.113	.087	.160	.116	.140	.094	.138
Asymptotische Signifikanz (2-seitig)		.000 ^c	.063 ^c	.200 ^{c,d}	.001 ^c	.049 ^c	.006 ^c	.200 ^{c,d}	.008 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 270: Zielorientierungen – Mittelwertsvergleich – Nichtparametrischer Test

		Statistik für Test ^a	
		yBestrebungen im Studium - xBestrebungen im Studium	zBestrebungen im Studium - xBestrebungen im Studium
Z		-.812 ^b	-1.262 ^b
Asymptotische Signifikanz (2-seitig)		.417	.207

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 271: Zielorientierungen – Mittelwertsvergleich – Parametrischer Test

Test bei gepaarten Stichproben

Gepaarte Differenzen

T | df | Sig. (2-seitig)

10. Anhang

		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz				
					Untere	Obere			
Paaren 1	xBestrebungen im Studium - y-Bestrebungen im Studium	-.02636	.30211	.03967	-.10580	.05307	-.665	57	.509
Paaren 2	xBestrebungen im Studium - zBestrebungen im Studium	-.03233	.24368	.03200	-.09640	.03175	-1.010	57	.317

Tabelle 272: Fachspezifische Überzeugung – Mittelwertsvergleich – Nichtparametrischer Test

Statistik für Test ^a						
	yChemie als Wissenschaft - xChemie als Wissenschaft	zChemie als Wissenschaft - xChemie als Wissenschaft	yUnterrichten von Chemie in der Schule - xUnterrichten von Chemie in der Schule	zUnterrichten von Chemie in der Schule - xUnterrichten von Chemie in der Schule	yLernen von Chemie in der Schule - xLernen von Chemie in der Schule	zLernen von Chemie in der Schule - xLernen von Chemie in der Schule
Z	-1.973 ^b	-3.590 ^b	-2.140 ^b	-2.820 ^b	-4.044 ^b	-1.710 ^b
Asymptotische Signifikanz (2-seitig)	.048	.000	.032	.005	.000	.087

a. Wilcoxon-Test

b. Basiert auf positiven Rängen.

Tabelle 273: Fachspezifische Überzeugung – Mittelwertsvergleich – Parametrischer Test

Test bei gepaarten Stichproben									
		Gepaarte Differenzen							
	Mittelwert	Std.-Abwei- chung	Standardfeh- ler des Mittel- wertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)	
				Untere	Obere				
Paaren 1	xChemie als Wissen- schaft - yChemie als Wis- senschaft	.08103	.29347	.03853	.00387	.15820	2.103	57	.040
Paaren 2	xChemie als Wissen- schaft - zChemie als Wis- senschaft	.12414	.23791	.03124	.06158	.18669	3.974	57	.000
Paaren 3	xUnterrichten von Che- mie in der Schule - yUn- terrachten von Chemie in der Schule	.06650	.24394	.03203	.00236	.13064	2.076	57	.042
Paaren 4	xUnterrichten von Che- mie in der Schule - zUn- terrachten von Chemie in der Schule	.06814	.19928	.02617	.01575	.12054	2.604	57	.012
Paaren 5	xLernen von Chemie in der Schule - yLernen von Chemie in der Schule	.23491	.37615	.04939	.13601	.33382	4.756	57	.000
Paaren 6	xLernen von Chemie in der Schule - zLernen von Chemie in der Schule	.10499	.41127	.05400	-.00315	.21312	1.944	57	.057

Tabelle 274: Selbstwirksamkeitserwartung – Mittelwertsvergleich – Nichtparametrischer Test

Statistik für Test ^a								
	yPlanung von Experimenten - xPlanung von Experimenten	zPlanung von Experimenten - xPlanung von Experimenten	yDurchführung von Experimenten - xDurchführung von Experimenten	zDurchführung von Experimenten - xDurchführung von Experimenten	ydidaktische Rekonstruktion unterrichtsrelevanter Inhalte - xdidaktische Rekonstruktion unterrichtsrelevanter Inhalte	zdidaktische Rekonstruktion unterrichtsrelevanter Inhalte - xdidaktische Rekonstruktion unterrichtsrelevanter Inhalte	yUmsetzung der didaktischen Rekonstruktion im Unterricht - xUmsetzung der didaktischen Rekonstruktion im Unterricht	zUmsetzung der didaktischen Rekonstruktion im Unterricht - xUmsetzung der didaktischen Rekonstruktion im Unterricht
Z	-3.409 ^b	-4.766 ^b	-4.146 ^b	-3.832 ^b	-2.332 ^b	-3.674 ^b	-2.620 ^b	-3.525 ^b
Asymptotische Signifikanz (2-seitig)	.001	.000	.000	.000	.020	.000	.009	.000

a. Wilcoxon-Test

b. Basiert auf positiven Rängen.

Tabelle 275: Selbstwirksamkeitserwartung – Mittelwertsvergleich – Parametrischer Test

Test bei gepaarten Stichproben									
Gepaarte Differenzen									
		Mittelwert	Std.-Abwei- chung	Standardfeh- ler des Mittel- wertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	xPlanung von Experimen- ten - yPlanung von Experimenten	.26316	.57837	.07661	.10970	.41662	3.435	56	.001

Paaren 2	xPlanung von Experimenten - zPlanung von Experimenten	.32069	.45567	.05983	.20088	.44050	5.360	57	.000
Paaren 3	xDurchführung von Experimenten - yDurchführung von Experimenten	.30702	.47485	.06290	.18102	.43301	4.881	56	.000
Paaren 4	xDurchführung von Experimenten - zDurchführung von Experimenten	.30172	.52515	.06896	.16364	.43981	4.376	57	.000
Paaren 5	xdidaktische Rekonstruktion unterrichtsrelevanter Inhalte - ydidaktische Rekonstruktion unterrichtsrelevanter Inhalte	.12284	.38722	.05084	.02103	.22466	2.416	57	.019
Paaren 6	xdidaktische Rekonstruktion unterrichtsrelevanter Inhalte - zdidaktische Rekonstruktion unterrichtsrelevanter Inhalte	.20690	.37940	.04982	.10714	.30665	4.153	57	.000
Paaren 7	xUmsetzung der didaktischen Rekonstruktion im Unterricht - yUmsetzung der didaktischen Rekonstruktion im Unterricht	.17241	.47896	.06289	.04648	.29835	2.742	57	.008
Paaren 8	xUmsetzung der didaktischen Rekonstruktion im Unterricht - zUmsetzung der didaktischen Rekonstruktion im Unterricht	.22414	.47988	.06301	.09796	.35032	3.557	57	.001

Tabelle 276: Standardisierte Residuen – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest									
		SR Zielorientierung	SR FÜ Chemie als Wissenschaft	SR FÜ Unterrichten von Chemie	SR FÜ Lernen von Chemie	SR SE Planung von Experimenten	SR SE Durchführung von Experimenten	SR SE Didaktische Rekonstruktion	SR SE Umsetzung der didaktischen Rekonstruktion
N		58	58	58	58	58	58	58	58
Parameter der Normalverteilung ^{a,b}	Mittelwert	.0000000	.0000000	.0000000	.0000000	.0000000	.0000000	.0000000	.0000000
	Std.-Abweichung	.99118926	.99118926	.99118926	.99118926	.99118926	.99118926	.99118926	.99118926
Extremste Differenzen	Absolut	.064	.068	.081	.092	.085	.088	.096	.095
	Positiv	.064	.068	.081	.092	.064	.088	.078	.073
	Negativ	-.054	-.045	-.066	-.089	-.085	-.068	-.096	-.095
Statistik für Test		.064	.068	.081	.092	.085	.088	.096	.095
Asymptotische Signifikanz (2-seitig)		.200 ^{c,d}	.200 ^{c,d}	.200 ^{c,d}	.200 ^{c,d}	.200 ^{c,d}	.200 ^{c,d}	.200 ^{c,d}	.200 ^{c,d}

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

d. Dies ist eine untere Grenze der echten Signifikanz.

Kolmogorov-Smirnov-Anpassungstest							
		SR Reflexionskompetenztest	SR Reflexionstiefe	SR Reflexionsbreite	SR Perspektiven	SR Literatur	SR Gesamtpunktzahl
N		58	58	58	57	57	58
Parameter der Normalverteilung ^{a,b}	Mittelwert	.0000000	.0000000	.0000000	.0000000	.0000000	.0000000
	Std.-Abweichung	.99118926	.99118926	.99118926	.99103121	.99103121	.99118926
Extremste Differenzen	Absolut	.100	.162	.090	.108	.131	.151
	Positiv	.079	.162	.061	.108	.131	.151
	Negativ	-.100	-.082	-.090	-.049	-.064	-.064
Statistik für Test		.100	.162	.090	.108	.131	.151
Asymptotische Signifikanz (2-seitig)		.200 ^{c,d}	.001 ^c	.200 ^{c,d}	.095 ^c	.017 ^c	.002 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 277: Korrelationen – Parametrisch

Korrelationen														
		SR Zielorientierung	SR FÜ Chemie als Wissenschaft	SR FÜ Unterrichten von Chemie	SR FÜ Lernen von Chemie	SR SE Planung von Experimenten	SR SE Durchführung von Experimenten	SR SE Didaktische Rekonstruktion	SR SE Umsetzung der didaktischen Rekonstruktion	SR Reflexionskompetenztest	SR Reflexionstiefe	SR Reflexionsbreite	SR Perspektiven	SR Gesamtpunktzahl
SR Zielorientierung	Korrelation nach Pearson	1	.375**	.162	.265*	.105	.086	.027	.159	.116	-.234	-.254	-.128	-.189

10. Anhang

	Signifikanz (2-seitig)		.004	.223	.045	.435	.522	.842	.232	.385	.077	.054	.345	.159	.067
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR FÜ Chemie als Wissenschaft	Korrelation nach Pearson	.375**	1	.342**	.364**	.094	.114	.114	-.026	-.141	-.224	-.351**	-.158	-.058	-.244
	Signifikanz (2-seitig)	.004		.009	.005	.485	.392	.392	.844	.291	.092	.007	.241	.669	.065
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR FÜ Unterrichten von Chemie	Korrelation nach Pearson	.162	.342**	1	.243	.130	.099	.002	-.132	.053	-.335*	-.164	-.186	-.155	-.338**
	Signifikanz (2-seitig)	.223	.009		.066	.329	.460	.987	.324	.694	.010	.219	.167	.251	.009
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR FÜ Lernen von Chemie	Korrelation nach Pearson	.265*	.364**	.243	1	.203	.093	.022	-.040	-.303*	-.196	-.148	-.115	-.097	-.204
	Signifikanz (2-seitig)	.045	.005	.066		.126	.487	.869	.766	.021	.141	.267	.393	.471	.125
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR SE Planung von Experimenten	Korrelation nach Pearson	.105	.094	.130	.203	1	.672**	.400**	.447**	.142	.088	.185	.093	.162	.091
	Signifikanz (2-seitig)	.435	.485	.329	.126		.000	.002	.000	.289	.512	.164	.490	.228	.497
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR SE Durchführung von Experimenten	Korrelation nach Pearson	.086	.114	.099	.093	.672**	1	.471**	.504**	.176	.184	.135	.198	.241	.179
	Signifikanz (2-seitig)	.522	.392	.460	.487	.000		.000	.000	.186	.168	.312	.140	.071	.178
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR SE Didaktische Rekonstruktion	Korrelation nach Pearson	.027	.114	.002	.022	.400**	.471**	1	.544**	.010	-.180	-.060	-.132	-.021	-.191
	Signifikanz (2-seitig)	.842	.392	.987	.869	.002	.000		.000	.942	.177	.654	.326	.876	.151
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR SE Umsetzung der didak Rekonstruktion	Korrelation nach Pearson	.159	-.026	-.132	-.040	.447**	.504**	.544**	1	.312*	.168	.121	.113	.168	.165
	Signifikanz (2-seitig)	.232	.844	.324	.766	.000	.000	.000		.017	.207	.367	.402	.213	.216
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR Reflexionskompetenztest	Korrelation nach Pearson	.116	-.141	.053	-.303*	.142	.176	.010	.312*	1	.183	.063	.154	.179	.183
	Signifikanz (2-seitig)	.385	.291	.694	.021	.289	.186	.942	.017		.169	.639	.252	.182	.170
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR Reflexionsstiefe	Korrelation nach Pearson	-.234	-.224	-.335*	-.196	.088	.184	-.180	.168	.183	1	.430**	.674**	.727*	.998**
	Signifikanz (2-seitig)	.077	.092	.010	.141	.512	.168	.177	.207	.169		.001	.000	.000	.000
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR Reflexionsbreite	Korrelation nach Pearson	-.254	-.351**	-.164	-.148	.185	.135	-.060	.121	.063	.430**	1	.474**	.442*	.479**
	Signifikanz (2-seitig)	.054	.007	.219	.267	.164	.312	.654	.367	.639	.001		.000	.001	.000
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58
SR Perspektiven	Korrelation nach Pearson	-.128	-.158	-.186	-.115	.093	.198	-.132	.113	.154	.674**	.474**	1	.915*	.676**
	Signifikanz (2-seitig)	.345	.241	.167	.393	.490	.140	.326	.402	.252	.000	.000		.000	.000
	N	57	57	57	57	57	57	57	57	57	57	57	57	57	57
SR Literatur	Korrelation nach Pearson	-.189	-.058	-.155	-.097	.162	.241	-.021	.168	.179	.727**	.442**	.915**	1	.725**
	Signifikanz (2-seitig)	.159	.669	.251	.471	.228	.071	.876	.213	.182	.000	.001	.000		.000
	N	57	57	57	57	57	57	57	57	57	57	57	57	57	57
SR Gesamtpunktzahl	Korrelation nach Pearson	-.243	-.244	-.338**	-.204	.091	.179	-.191	.165	.183	.998**	.479**	.676**	.725*	1
	Signifikanz (2-seitig)	.067	.065	.009	.125	.497	.178	.151	.216	.170	.000	.000	.000	.000	
	N	58	58	58	58	58	58	58	58	58	58	58	57	57	58

** Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant.

* Die Korrelation ist auf dem Niveau von 0.05 (2-seitig) signifikant.

Tabelle 278: Korrelationen – Nichtparametrisch

		SR Zielori- entierung	SR FÜ Chemie als Wissen- schaft	SR FÜ Unter- richten von Chemie	SR FÜ Lernen von Chemie	SR SE Plan- nung von Experi- menten	SR SE Durch- führung von Experi- menten	SR SE Didak- tische Rekon- struktion	SR SE Um- setzung der didak Rekon- struktion	SR Re- flexi- onskom- petenztest	SR Re- flexi- ons- tiefe	SR Re- flexi- ons- breite	SR Per- spek- tiven	SR Lite- ratur	SR Ge- samt punk- tzahl
Spe- arman- Rho	SR Zielori- entierung	Korrelati- onskoeffi- zient	1.000	.354**	.187	.240	.076	.135	.001	.185	.117	-.252	-.210	-.132	-
		Sig. (2- seitig)	.	.006	.160	.070	.569	.314	.997	.164	.382	.056	.114	.329	.178
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR FÜ Chemie als Wissen- schaft	Korrelati- onskoeffi- zient	.354**	1.000	.253	.369**	.061	.152	.096	-.015	-.147	-.187	-.342**	-.165	-.036	-.223
		Sig. (2- seitig)	.006	.	.056	.004	.648	.255	.476	.911	.270	.160	.009	.221	.092
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR FÜ Unter- richten von Chemie	Korrelati- onskoeffi- zient	.187	.253	1.000	.143	.210	.121	.154	.007	.081	-.340**	-.122	-.158	-.136	-.350**
		Sig. (2- seitig)	.160	.056	.	.283	.113	.365	.248	.958	.543	.009	.362	.241	.007
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR FÜ Lernen von Chemie	Korrelati- onskoeffi- zient	.240	.369**	.143	1.000	.109	.121	.049	-.121	-.325*	-.212	-.209	-.088	-.116	-.252
		Sig. (2- seitig)	.070	.004	.283	.	.416	.364	.714	.367	.013	.110	.115	.514	.056
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR SE Plan- nung von Experi- menten	Korrelati- onskoeffi- zient	.076	.061	.210	.109	1.000	.632**	.433**	.435**	.125	-.005	.268*	.024	.063	.001
		Sig. (2- seitig)	.569	.648	.113	.416	.	.000	.001	.001	.349	.973	.042	.860	.995
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR SE Durch- führung von Experi- menten	Korrelati- onskoeffi- zient	.135	.152	.121	.121	.632**	1.000	.445**	.545**	.195	.168	.170	.167	.196	.154
		Sig. (2- seitig)	.314	.255	.365	.364	.000	.	.000	.000	.142	.206	.203	.214	.248
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR SE Didak- tische Rekon- struktion	Korrelati- onskoeffi- zient	.001	.096	.154	.049	.433**	.445**	1.000	.465**	-.017	-.242	.049	-.217	-.115	-.254
		Sig. (2- seitig)	.997	.476	.248	.714	.001	.000	.	.000	.898	.067	.715	.104	.054
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR SE Um- setzung der didak Re- konstruktion	Korrelati- onskoeffi- zient	.185	-.015	.007	-.121	.435**	.545**	.465**	1.000	.341**	.138	.282*	.055	.137	.142
		Sig. (2- seitig)	.164	.911	.958	.367	.001	.000	.000	.	.009	.301	.032	.686	.286
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR Reflexi- onskompe- tenztest	Korrelati- onskoeffi- zient	.117	-.147	.081	-.325*	.125	.195	-.017	.341**	1.000	.118	.057	.090	.123	.120
		Sig. (2- seitig)	.382	.270	.543	.013	.349	.142	.898	.009	.	.380	.673	.363	.369
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR Reflexi- ons- tiefe	Korrelati- onskoeffi- zient	-.252	-.187	-.340**	-.212	-.005	.168	-.242	.138	.118	1.000	.472**	.664**	.688**	.993**
		Sig. (2- seitig)	.056	.160	.009	.110	.973	.206	.067	.301	.380	.	.000	.000	.000
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR Reflexi- onsbreite	Korrelati- onskoeffi- zient	-.210	-.342**	-.122	-.209	.268*	.170	.049	.282*	.057	.472**	1.000	.459**	.436**	.533**
		Sig. (2- seitig)	.114	.009	.362	.115	.042	.203	.715	.032	.673	.000	.	.000	.000
		N	58	58	58	58	58	58	58	58	58	58	57	57	58
SR Per- spektiven	Korrelati- onskoeffi- zient	-.132	-.165	-.158	-.088	.024	.167	-.217	.055	.090	.664**	.459**	1.000	.928**	.653**
		Sig. (2- seitig)	.329	.221	.241	.514	.860	.214	.104	.686	.504	.000	.000	.	.000
		N	57	57	57	57	57	57	57	57	57	57	57	57	57
SR Litera- tur	Korrelati- onskoeffi- zient	-.178	-.036	-.136	-.116	.063	.196	-.115	.137	.123	.688**	.436**	.928**	1.000	.674**
		Sig. (2- seitig)	.186	.788	.314	.390	.640	.145	.392	.310	.363	.000	.001	.	.000
		N	57	57	57	57	57	57	57	57	57	57	57	57	57

10. Anhang

Sig. (2-seitig)	.047	.092	.007	.056	.995	.248	.054	.286	.369	.000	.000	.000	.000	.
N	58	58	58	58	58	58	58	58	58	58	58	57	57	58

10.3.16. Statistiken zum Kapitel 7.2.3

Tabelle 279: Wissenstest – SuS – Test auf Normalverteilung MZP1

Kolmogorov-Smirnov-Anpassungstest		Average Percentage Pre	Average Percentage Post	Average Percentage FollowUp
N		170	170	170
Parameter der Normalverteilung ^{a,b}	Mittelwert	30.7199	47.0287	36.7196
	Std.-Abweichung	10.54723	14.00826	14.97761
Extremste Differenzen	Absolut	.097	.101	.075
	Positiv	.076	.084	.075
	Negativ	-.097	-.101	-.068
Statistik für Test		.097	.101	.075
Asymptotische Signifikanz (2-seitig)		.000 ^c	.000 ^c	.019 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

Tabelle 280: Wissenstest – SuS – Test auf Normalverteilung MZP2

Kolmogorov-Smirnov-Anpassungstest		Average Percentage Pre	Average Percentage Post	Average Percentage FollowUp
N		171	171	171
Parameter der Normalverteilung ^{a,b}	Mittelwert	30.7846	51.2518	40.5533
	Std.-Abweichung	12.85386	14.34911	14.84742
Extremste Differenzen	Absolut	.092	.071	.090
	Positiv	.092	.071	.090
	Negativ	-.052	-.053	-.044
Statistik für Test		.092	.071	.090
Asymptotische Signifikanz (2-seitig)		.001 ^c	.035 ^c	.002 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

Tabelle 281: Wissenstest – SuS – Test auf Normalverteilung MZP3

Kolmogorov-Smirnov-Anpassungstest		Average Percentage Pre	Average Percentage Post	Average Percentage FollowUp
N		159	159	159
Parameter der Normalverteilung ^{a,b}	Mittelwert	34.1552	53.5264	41.4649
	Std.-Abweichung	11.59130	16.14432	14.17833
Extremste Differenzen	Absolut	.099	.093	.098
	Positiv	.099	.054	.098
	Negativ	-.092	-.093	-.068
Statistik für Test		.099	.093	.098
Asymptotische Signifikanz (2-seitig)		.001 ^c	.002 ^c	.001 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

Tabelle 282: Wissenstest – SuS – Mittelwertvergleiche – Nichtparametrisch – MZP1

Statistik für Test ^a		
	Average Percentage Post - Average Percentage Pre	Average Percentage FollowUp - Average Percentage Pre
Z	-9.506 ^b	-4.580 ^b
Asymptotische Signifikanz (2-seitig)	.000	.000

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 283: Wissenstest – SuS – Mittelwertvergleiche – Nichtparametrisch – MZP2

Statistik für Test ^a		
	Average Percentage Post - Average Percentage Pre	Average Percentage FollowUp - Average Percentage Pre
Z	-10.334 ^b	-7.073 ^b
Asymptotische Signifikanz (2-seitig)	.000	.000

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 284: Wissenstest – SuS – Mittelwertvergleiche – Nichtparametrisch - MZP3

Statistik für Test^a

	Average Percentage Post - Average Percentage Pre	Average Percentage FollowUp - Average Percentage Pre
Z	-9.884 ^b	-5.581 ^b
Asymptotische Signifikanz (2-seitig)	.000	.000

a. Wilcoxon-Test
b. Basiert auf negativen Rängen.

Tabelle 285: Wissenstest – SuS – Mittelwertvergleiche – Parametrisch – MZP1

Test bei gepaarten Stichproben									
		Gepaarte Differenzen							
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	Average Percentage Pre - Average Percentage Post	-16.30878	15.97649	1.22534	-18.72773	-13.88984	-13.310	169	.000
Paaren 2	Average Percentage Pre - Average Percentage FollowUp	-5.99967	15.57847	1.19481	-8.35835	-3.64099	-5.021	169	.000

Tabelle 286: Wissenstest – SuS – Mittelwertvergleiche – Parametrisch – MZP2

Test bei gepaarten Stichproben									
		Gepaarte Differenzen							
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	Average Percentage Pre - Average Percentage Post	-20.46718	16.14120	1.23435	-22.90381	-18.03056	-16.581	170	.000
Paaren 2	Average Percentage Pre - Average Percentage FollowUp	-9.76870	15.71744	1.20194	-12.14136	-7.39605	-8.127	170	.000

Tabelle 287: Wissenstest – SuS – Mittelwertvergleiche – Parametrisch – MZP3

Test bei gepaarten Stichproben									
		Gepaarte Differenzen							
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	Average Percentage Pre - Average Percentage Post	-19.37125	16.42437	1.30254	-21.94388	-16.79862	-14.872	158	.000
Paaren 2	Average Percentage Pre - Average Percentage FollowUp	-7.30969	14.93957	1.18478	-9.64975	-4.96963	-6.170	158	.000

Tabelle 288: Residualer Lernzuwachs – SuS – Berechnung von Unterschieden zwischen den drei Messzeitpunkten

Einfaktorielle ANOVA					
SR CK SuS					
	Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
Zwischen den Gruppen	12.345	2	6.173	6.317	.002
Innerhalb der Gruppen	485.655	497	.977		
Gesamt	498.000	499			

Tabelle 289: Residualer Lernzuwachs – SuS – Post-Hoc Test

Mehrfachvergleiche							
Abhängige Variable: SR CK SuS							
	(I) Messzeitpunkt	(J) Messzeitpunkt	Mittlere Differenz (I-J)	Std.-Fehler	Signifikanz	95%-Konfidenzintervall Untergrenze	Obergrenze
Tukey-HSD	1.00	2.00	-.29103790*	.10706322	.019	-.5427172	-.0393585
		3.00	-.36333116*	.10905876	.003	-.6197015	-.1069608
	2.00	1.00	.29103790*	.10706322	.019	.0393585	.5427172

10. Anhang

		3.00		-.07229326	.10890454	.785	-.3283011	.1837146
	3.00	1.00		.36333116*	.10905876	.003	.1069608	.6197015
		2.00		.07229326	.10890454	.785	-.1837146	.3283011
Bonferroni	1.00	2.00		-.29103790*	.10706322	.020	-.5482157	-.0338601
		3.00		-.36333116*	.10905876	.003	-.6253025	-.1013598
	2.00	1.00		.29103790*	.10706322	.020	.0338601	.5482157
		3.00		-.07229326	.10890454	1.000	-.3338941	.1893076
	3.00	1.00		.36333116*	.10905876	.003	.1013598	.6253025
		2.00		.07229326	.10890454	1.000	-.1893076	.3338941

*. Die Differenz der Mittelwerte ist auf dem Niveau 0.05 signifikant.

10.3.17. Statistiken zum Kapitel 7.2.4

Tabelle 290: Reflexionskompetenztest – Residualer Lernzuwachs – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest		Standardized Residual
N		58
Parameter der Normalverteilung ^{a,b}	Mittelwert	.0000000
	Std.-Abweichung	.99118926
Extremste Differenzen	Absolut	.100
	Positiv	.079
	Negativ	-.100
Statistik für Test		.100
Asymptotische Signifikanz (2-seitig)		.200 ^{c,d}

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 291: Reflexionskompetenztest – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
Standardized Residual	Feedback	20	.2933465	.76272018	.17054942
	Kein Feedback	21	-.3793577	.98844076	.21569546

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz Untere Obere
Standardized Residual	Varianzen sind gleich	1.441	.237	2.431	39	.020	.67270417	.27672530	.11297442 1.2324339
	Varianzen sind nicht gleich			2.446	37.426	.019	.67270417	.27497570	.11576470 1.2296436

Tabelle 292: Reflexionskompetenztest – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
Standardized Residual	Feedback	20	.2933465	.76272018	.17054942
	Unterrichtsvideos	17	.1235048	1.12927214	.27388872

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz

								Standard- fehler der Differenz	Untere	Obere
Standardized Residual	Varianzen sind gleich	2.895	.098	.543	35	.591	.16984166	.31274281	-.46505999	.80474332
	Varianzen sind nicht gleich			.526	27.351	.603	.16984166	.32264863	-.49178165	.83146497

Tabelle 293: Kodiermanual – Residualer Lernzuwachs – Test auf Normalverteilung

		Kolmogorov-Smirnov-Anpassungstest				
		SR Gesamtpunktzahl Kodiermanual	SR Reflexionsbreite	SR Reflexionstiefe	SR Perspektiven	SR Literatur
N		58	58	58	57	57
Parameter der Normalverteilung ^{a,b}	Mittelwert	.0000000	.0000000	.0000000	.0000000	.0000000
	Std.-Abweichung	.99118926	.99118926	.99118926	.99103121	.99103121
Extremste Differenzen	Absolut	.151	.090	.162	.108	.131
	Positiv	.151	.061	.162	.108	.131
	Negativ	-.064	-.090	-.082	-.049	-.064
Statistik für Test		.151	.090	.162	.108	.131
Asymptotische Signifikanz (2-seitig)		.002 ^c	.200 ^{c,d}	.001 ^c	.095 ^c	.017 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 294: Kodiermanual – Gesamtpunktzahl – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. IG2

Statistik für Test ^a		SR Gesamtpunktzahl Kodiermanual
Mann-Whitney-U		201.000
Wilcoxon-W		411.000
Z		-.235
Asymptotische Signifikanz (2-seitig)		.814

a. Gruppenvariable: gruppe Gruppenzuordnung

Tabelle 295: Kodiermanual – Gesamtpunktzahl – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG

Statistik für Test ^a		SR Gesamtpunktzahl Kodiermanual
Mann-Whitney-U		146.000
Wilcoxon-W		299.000
Z		-.731
Asymptotische Signifikanz (2-seitig)		.465
Exakte Signifikanz [2*(1-seitige Sig.)]		.478 ^b

a. Gruppenvariable: gruppe Gruppenzuordnung

b. Nicht für Bindungen korrigiert.

Tabelle 296: Kodiermanual – Gesamtpunktzahl – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Gesamtpunktzahl Kodiermanual	Feedback	20	.0351702	1.01636786	.22726676
	Kein Feedback	21	.0690540	1.01416608	.22130918

Test bei unabhängigen Stichproben								
Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit				
F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
							Untere	Obere

10. Anhang

SR Gesamtpunktzahl Kodiermanual	Varianzen sind gleich	.139	.711	-.107	39	.915	-	.31720186	-	.60771749
	Varianzen sind nicht gleich			-.107	38.894	.915	.03388382	.31721907	.67548514	.60780824
							.03388382		.67557589	

Tabelle 297: Kodiermanual – Gesamtpunktzahl – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Gesamtpunktzahl Kodiermanual	Feedback	20	.0351702	1.01636786	.22726676
	Unterrichtsvideos	17	-.1266787	.98121706	.23798009

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
SR Gesamtpunktzahl Kodiermanual	Varianzen sind gleich	.009	.925	.490	35	.627	.16184883	.33003328	-.50815435	.83185201
	Varianzen sind nicht gleich			.492	34.399	.626	.16184883	.32906642	-.50660931	.83030697

Tabelle 298: Kodiermanual – Reflexionsbreite – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Reflexionsbreite	Feedback	20	.1571165	.79746890	.17831947
	Kein Feedback	21	.0513574	1.02721961	.22415770

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
SR Reflexionsbreite	Varianzen sind gleich	2.282	.139	.367	39	.716	.10575906	.28821563	-.47721208	.68873020
	Varianzen sind nicht gleich			.369	37.510	.714	.10575906	.28643412	-.47434522	.68586335

Tabelle 299: Kodiermanual – Reflexionsbreite – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Reflexionsbreite	Feedback	20	.1571165	.79746890	.17831947
	Unterrichtsvideos	17	-.2482845	1.15152486	.27928580

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
SR Reflexionsbreite	Varianzen sind gleich	.627	.434	1.260	35	.216	.40540097	.32176968	-.24782621	1.05862815
	Varianzen sind nicht gleich			1.223	27.812	.231	.40540097	.33135840	-.27356276	1.08436470

Tabelle 300: Kodiermanual – Reflexionstiefe – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. IG2

Statistik für Test ^a		SR Reflexionstiefe
Mann-Whitney-U		200.000
Wilcoxon-W		410.000
Z		-.261
Asymptotische Signifikanz (2-seitig)		.794

a. Gruppenvariable: gruppe Gruppenzuordnung

Tabelle 301: Kodiermanual – Reflexionstiefe – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG

Statistik für Test ^a	
SR Reflexionstiefe	
Mann-Whitney-U	146.000
Wilcoxon-W	299.000
Z	-.731
Asymptotische Signifikanz (2-seitig)	.465
Exakte Signifikanz [2*(1-seitige Sig.)]	.478 ^b

a. Gruppenvariable: gruppe Gruppenzuordnung
b. Nicht für Bindungen korrigiert.

Tabelle 302: Kodiermanual – Reflexionstiefe – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken									
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes				
SR Reflexionstiefe	Feedback	20	.0265407	1.03236379	.23084356				
	Kein Feedback	21	.0676761	1.00084760	.21840285				

Test bei unabhängigen Stichproben									
Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
SR Reflexionstiefe	Varianzen sind gleich	.170	.682	-.130	39	.898	-.04113540	.31754067	Untere: -.68342202 Obere: .60115123
	Varianzen sind nicht gleich			-.129	38.746	.898	-.04113540	.31778697	Untere: -.68405511 Obere: .60178431

Tabelle 303: Kodiermanual – Reflexionstiefe – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken									
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes				
SR Reflexionstiefe	Feedback	20	.0265407	1.03236379	.23084356				
	Unterrichtsvideos	17	-.1148243	.98049569	.23780514				

Test bei unabhängigen Stichproben									
Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
SR Reflexionstiefe	Varianzen sind gleich	.004	.949	.425	35	.674	.14136498	.33284765	Untere: -.53435167 Obere: .81708163
	Varianzen sind nicht gleich			.427	34.537	.672	.14136498	.33142123	Untere: -.53177880 Obere: .81450876

PCK-Test

Tabelle 304: PCK-Test – Residualer Lernzuwachs – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest		
		Standardized Residual
N		58
Parameter der Normalverteilung ^{a,b}	Mittelwert	.0000000
	Std.-Abweichung	.99118926
Extremste Differenzen	Absolut	.128
	Positiv	.089
	Negativ	-.128
Statistik für Test		.128
Asymptotische Signifikanz (2-seitig)		.019 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.
b. Aus den Daten berechnet.
c. Signifikanzkorrektur nach Lilliefors.

Tabelle 305: PCK-Test – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. IG2

Statistik für Test^a

10. Anhang

	Standardized Residual
Mann-Whitney-U	200.500
Wilcoxon-W	431.500
Z	-.248
Asymptotische Signifikanz (2-seitig)	.804

a. Gruppenvariable: gruppe Gruppenzuordnung

Tabelle 306: PCK-Test – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG

Statistik für Test ^a	Standardized Residual
Mann-Whitney-U	127.500
Wilcoxon-W	337.500
Z	-1.296
Asymptotische Signifikanz (2-seitig)	.195
Exakte Signifikanz [2*(1-seitige Sig.)]	.198 ^b

a. Gruppenvariable: gruppe Gruppenzuordnung

b. Nicht für Bindungen korrigiert.

Tabelle 307: PCK-Test – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
Standardized Residual	Feedback	20	-.1263001	1.06730061	.23865567
	Kein Feedback	21	-.1832867	1.09965364	.23996410

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
Standardized Residual	Varianzen sind gleich	.003	.954	.168	39	.867	.05698659	.33868939	-.62807736	.74205053
	Varianzen sind nicht gleich			.168	38.984	.867	.05698659	.33843655	-.62757486	.74154803

Tabelle 308: PCK-Test – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
Standardized Residual	Feedback	20	-.1263001	1.06730061	.23865567
	Unterrichtsvideos	17	.3750014	.64983669	.15760855

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
Standardized Residual	Varianzen sind gleich	5.417	.026	-1.687	35	.101	-.50130159	.29715785	1.10456408	-.10196091
	Varianzen sind nicht gleich			-1.753	31.967	.089	-.50130159	.28600172	1.08389195	-.08128878

CK-Test

Tabelle 309: CK-Test – Residualer Lernzuwachs – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest			SR CK-Test
N			58
Parameter der Normalverteilung ^{a,b}	Mittelwert		.0000000
	Std.-Abweichung		.99118926
Extremste Differenzen	Absolut		.129
	Positiv		.129
	Negativ		-.091
Statistik für Test			.129
Asymptotische Signifikanz (2-seitig)			.018 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

Tabelle 310: CK-Test – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. IG2

Statistik für Test ^a	
SR CK-Test	
Mann-Whitney-U	174.500
Wilcoxon-W	384.500
Z	-.926
Asymptotische Signifikanz (2-seitig)	.354

a. Gruppenvariable: gruppe Gruppenzuordnung

Tabelle 311: CK-Test – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG

Statistik für Test ^a	
SR CK-Test	
Mann-Whitney-U	155.000
Wilcoxon-W	365.000
Z	-.457
Asymptotische Signifikanz (2-seitig)	.647
Exakte Signifikanz [2*(1-seitige Sig.)]	.662 ^b

a. Gruppenvariable: gruppe Gruppenzuordnung

b. Nicht für Bindungen korrigiert.

Tabelle 312: CK-Test – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken									
gruppe Gruppenzuordnung		N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes				
SR CK-Test	Feedback	20	-.1481320	.92240047	.20625501				
	Kein Feedback	21	.1588456	1.05477640	.23017108				

Test bei unabhängigen Stichproben									
Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	
SR CK-Test	Varianzen sind gleich	.322	.574	-.990	39	.328	-.30697768	.31009515	Untere: -.93420433 Obere: .32024898
	Varianzen sind nicht gleich			-.993	38.729	.327	-.30697768	.30906287	Untere: -.93225625 Obere: .31830090

Tabelle 313: CK-Test – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken									
gruppe Gruppenzuordnung		N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes				
SR CK-Test	Feedback	20	-.1481320	.92240047	.20625501				
	Unterrichtsvideos	17	-.0219481	1.01877326	.24708881				

Test bei unabhängigen Stichproben									
Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	
SR CK-Test	Varianzen sind gleich	.288	.595	-.395	35	.695	-.12618395	.31921180	Untere: -.77421836 Obere: .52185046
	Varianzen sind nicht gleich			-.392	32.697	.698	-.12618395	.32186023	Untere: -.78124401 Obere: .52887611

10. Anhang

Selbstwirksamkeitserwartung

Tabelle 314: Selbstwirksamkeitserwartung – Residualer Lernzuwachs – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest						
		SR Selbstwirksamkeit Gesamt	SR Planung von Experimenten	SR Durchführung von Experimenten	SR Didaktische Rekonstruktion	SR Umsetzung der didaktischen Rekonstruktion
N		17	16	16	17	17
Parameter der Normalverteilung ^{a,b}	Mittelwert	.1443031	.1774826	.1730604	.1358777	.1865772
	Std.-Abweichung	.84697804	1.04693393	.71726026	1.14641135	.82881143
Extremste Differenzen	Absolut	.122	.194	.150	.107	.085
	Positiv	.112	.194	.106	.107	.084
	Negativ	-.122	-.115	-.150	-.105	-.085
Statistik für Test		.122	.194	.150	.107	.085
Asymptotische Signifikanz (2-seitig)		.200 ^{c,d}	.108 ^c	.200 ^{c,d}	.200 ^{c,d}	.200 ^{c,d}

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 315: Selbstwirksamkeitserwartung – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken						
		gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Selbstwirksamkeit Gesamt	Feedback		20	-.2636498	1.10538540	.24717169
	Kein Feedback		21	.1342782	.97937543	.21371724

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
SR Selbstwirksamkeit Gesamt	Varianzen sind gleich	.163	.689	-1.221	39	.229	-.39792806	.32577199	1.05686410	-.26100798
	Varianzen sind nicht gleich			-1.218	37.903	.231	-.39792806	.32675511	1.05946469	-.26360857

Tabelle 316: Selbstwirksamkeitserwartung – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Selbstwirksamkeit Gesamt	Feedback	20	-.2636498	1.10538540	.24717169
	Unterrichtsvideos	17	.1443031	.84697804	.20542235

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
SR Selbstwirksamkeit Gesamt	Varianzen sind gleich	.784	.382	-1.242	35	.222	-.40795293	.32843704	1.07471556	-.25880971
	Varianzen sind nicht gleich			-1.269	34.670	.213	-.40795293	.32139102	1.06063352	-.24472767

Tabelle 317: Selbstwirksamkeitserwartung – Planung von Experimenten – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken						
		gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes

SR Planung von Experimenten		Feedback	20	-.1579150	1.15709569	.25873446
		Kein Feedback	21	.0151704	.77822957	.16982362

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
SR Planung von Experimenten	Varianzen sind gleich	1.498	.228	-.565	39	.576	.17308538	.30658283	-.79320770	.44703693
	Varianzen sind nicht gleich			-.559	33.067	.580	.17308538	.30948923	-.80269755	.45652679

Tabelle 318: Selbstwirksamkeitserwartung – Planung von Experimenten – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Planung von Experimenten	Feedback	20	-.1579150	1.15709569	.25873446
	Unterrichtsvideos	16	.1774826	1.04693393	.26173348

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
SR Planung von Experimenten	Varianzen sind gleich	.000	.987	-.901	34	.374	-.33539754	.37225287	1.09190639	.42111131
	Varianzen sind nicht gleich			-.911	33.434	.369	-.33539754	.36803252	1.08379600	.41300093

Tabelle 319: Selbstwirksamkeitserwartung – Durchführung von Experimenten – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Durchführung von Exerimenten	Feedback	20	-.3582733	1.07001327	.23926224
	Kein Feedback	21	.2093572	1.03833024	.22658223

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
SR Durchführung von Exerimenten	Varianzen sind gleich	.364	.550	-1.724	39	.093	-.56763050	.32927619	-1.23365445	.09839345
	Varianzen sind nicht gleich			-1.723	38.752	.093	-.56763050	.32952379	-1.23429192	.09903092

Tabelle 320: Selbstwirksamkeitserwartung – Durchführung von Experimenten – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Durchführung von Exerimenten	Feedback	20	-.3582733	1.07001327	.23926224
	Unterrichtsvideos	16	.1730604	.71726026	.17931506

Test bei unabhängigen Stichproben							
Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					
F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz

10. Anhang

								Standardfehler der Differenz	Untere	Obere
SR Durchführung von Experimenten	Varianzen sind gleich	4.829	.035	-1.702	34	.098	-.53133367	.31227060	-1.16594388	-.10327654
	Varianzen sind nicht gleich			-1.777	33.108	.085	-.53133367	.29899885	-1.13957624	-.07690890

Tabelle 321: Selbstwirksamkeitserwartung – Didaktische Rekonstruktion – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Didaktische Rekonstruktion	Feedback	20	-.2148169	.89944083	.20112108
	Kein Feedback	21	.0945913	.95331746	.20803093

Test bei unabhängigen Stichproben									
		Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit			
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
SR Didaktische Rekonstruktion	Varianzen sind gleich	.058	.811	-1.068	39	.292	-.30940817	.28977632	Untere: -.89553611 Obere: .27671977
	Varianzen sind nicht gleich			-1.069	38.997	.292	-.30940817	.28935542	Untere: -.89468598 Obere: .27586964

Tabelle 322: Selbstwirksamkeitserwartung – Didaktische Rekonstruktion – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Didaktische Rekonstruktion	Feedback	20	-.2148169	.89944083	.20112108
	Unterrichtsvideos	17	.1358777	1.14641135	.27804559

Test bei unabhängigen Stichproben									
		Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit			
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
SR Didaktische Rekonstruktion	Varianzen sind gleich	.578	.452	-1.042	35	.304	-.35069458	.33641261	Untere: 1.03364848 Obere: .33225932
	Varianzen sind nicht gleich			-1.022	30.168	.315	-.35069458	.34316037	Untere: 1.05135762 Obere: .34996845

Tabelle 323: Selbstwirksamkeitserwartung – Umsetzung der didaktischen Rekonstruktion – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Umsetzung der didaktischen Rekonstruktion	Feedback	20	-.0853854	1.13911613	.25471411
	Kein Feedback	21	-.0697193	.98780926	.21555765

Test bei unabhängigen Stichproben									
		Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit			
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
SR Umsetzung der didaktischen Rekonstruktion	Varianzen sind gleich	.446	.508	-.047	39	.963	-.01566608	.33250338	Untere: .68821765 Obere: .65688549
	Varianzen sind nicht gleich			-.047	37.626	.963	-.01566608	.33368305	Untere: .69139258 Obere: .66006043

Tabelle 324: Selbstwirksamkeitserwartung – Umsetzung der didaktischen Rekonstruktion – Residualer Lernzuwachs – Parametrisch – IG1 vs. KG

Gruppenstatistiken									
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes				
SR Umsetzung der didaktischen Rekonstruktion	Feedback	20	-.0853854	1.13911613	.25471411				
	Unterrichtsvideos	17	.1865772	.82881143	.20101630				

Test bei unabhängigen Stichproben									
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
SR Umsetzung der didaktischen Rekonstruktion	Varianzen sind gleich	2.134	.153	-.817	35	.419	-.27196260	.33290979	Untere: -.94780540 Obere: .40388021
	Varianzen sind nicht gleich			-.838	34.257	.408	-.27196260	.32447932	Untere: -.93120164 Obere: .38727644

Unterrichtsqualität

Tabelle 325: Unterrichtsqualität – Residualer Lernzuwachs – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest			SR CK SuS
N			500
Parameter der Normalverteilung ^{a,b}	Mittelwert		.0000000
	Std.-Abweichung		.99899749
Extremste Differenzen	Absolut		.033
	Positiv		.023
	Negativ		-.033
Statistik für Test			.033
Asymptotische Signifikanz (2-seitig)			.200 ^{c,d}

a. Die zu testende Verteilung ist eine Normalverteilung.
b. Aus den Daten berechnet.
c. Signifikanzkorrektur nach Lilliefors.
d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 326: Unterrichtsqualität – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2 – Gesamtstichprobe

Gruppenstatistiken									
	gruppe Gruppeneinteilung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes				
SR CK SuS	Feedback	255	-.1241154	.96171025	.06022463				
	Kein Feedback	245	.1291813	1.02241699	.06531983				

Test bei unabhängigen Stichproben									
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
SR CK SuS	Varianzen sind gleich	1.134	.287	-2.854	498	.004	-.25329674	.08873763	Untere: -.42764302 Obere: -.07895045
	Varianzen sind nicht gleich			-2.851	492.958	.005	-.25329674	.08884642	Untere: -.42786111 Obere: -.07873237

Tabelle 327: Unterrichtsqualität – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2 – MZP1

Gruppenstatistiken									
	gruppe Gruppeneinteilung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes				
SR CK SuS	Feedback	90	-.2885019	.97942293	.10324024				
	Kein Feedback	80	-.1324682	.93753196	.10481926				

Test bei unabhängigen Stichproben									
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
	Varianzen sind gleich								Untere: Obere:
	Varianzen sind nicht gleich								Untere: Obere:

10. Anhang

SR CK SuS	Varianzen sind gleich	.266	.606	-1.058	168	.292	-.15603378	.14750532	-.44723661	.13516905
	Varianzen sind nicht gleich			-1.061	167.064	.290	-.15603378	.14712452	-.44649663	.13442907

Tabelle 328: Unterrichtsqualität – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2 – MZP2

Gruppenstatistiken										
	gruppe	Gruppeneinteilung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes				
SR CK SuS		Feedback	81	.0195976	.93054806	.10339423				
		Kein Feedback	90	.1266931	.96786348	.10202177				

Test bei unabhängigen Stichproben										
Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
SR CK SuS	Varianzen sind gleich	.121	.728	-.736	169	.463	-.10709551	.14555680	-.39443925	.18024824
	Varianzen sind nicht gleich			-.737	168.251	.462	-.10709551	.14525429	-.39385126	.17966025

Tabelle 329: Unterrichtsqualität – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2 – MZP3

Gruppenstatistiken										
	gruppe	Gruppeneinteilung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes				
SR CK SuS		Feedback	84	-.0865673	.95721254	.10444045				
		Kein Feedback	75	.4112600	1.10775607	.12791265				

Test bei unabhängigen Stichproben										
Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
SR CK SuS	Varianzen sind gleich	2.322	.130	-3.040	157	.003	-.49782738	.16377585	-.82131566	-.17433909
	Varianzen sind nicht gleich			-3.015	147.220	.003	-.49782738	.16513466	-.82416794	-.17148681

Tabelle 330: Unterrichtsqualität – Residualer Lernzuwachs – Unterschied in den drei MZP – Feedbackgruppe

Tests der Zwischensubjekteffekte							
Abhängige Variable: SR CK SuS							
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	
Korrigiertes Modell	4,223 ^a	2	2,112	2,307	,102	,018	
Konstanter Term	3,573	1	3,573	3,903	,049	,015	
schülertag	4,223	2	2,112	2,307	,102	,018	
Fehler	230,698	252	,915				
Gesamt	238,849	255					
Korrigierte Gesamtvariation	234,921	254					

a. R-Quadrat = ,018 (korrigiertes R-Quadrat = ,010)

Tabelle 331: Unterrichtsqualität – Residualer Lernzuwachs – Unterschied in den drei MZP – Selbstreflexionsgruppe

Tests der Zwischensubjekteffekte							
Abhängige Variable: SR CK SuS							

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	11.445 ^a	2	5.723	5.685	.004	.045
Konstanter Term	4.450	1	4.450	4.421	.037	.018
Schülertag	11.445	2	5.723	5.685	.004	.045
Fehler	243.617	242	1.007			
Gesamt	259.151	245				
Korrigierte Gesamtvariation	255.062	244				

a. R-Quadrat = .045 (korrigiertes R-Quadrat = .037)

Tabelle 332: Unterrichtsqualität – Residualer Lernzuwachs – Unterschied in den drei MZP – Post-Hoc-Test – Selbstreflexionsgruppe

Multiple Comparisons							
Abhängige Variable: SR CK SuS							
Bonferroni							
(I)Schülertag	(J)Schülertag	Mittlere Differenz (I-J)	Standard Fehler	Sig.	95%-Konfidenzintervall		
1.2.....6	1.2.....6				Untergrenze	Obergrenze	
2. Schülertag	4. Schülertag	-.2591612	.15417161	.282	-.6308285	.1125061	
	6. Schülertag	-.5437282 [*]	.16126362	.003	-.9324925	-.1549639	
4. Schülertag	2. Schülertag	.2591612	.15417161	.282	-.1125061	.6308285	
	6. Schülertag	-.2845670	.15686869	.213	-.6627362	.0936023	
6. Schülertag	2. Schülertag	.5437282 [*]	.16126362	.003	.1549639	.9324925	
	4. Schülertag	.2845670	.15686869	.213	-.0936023	.6627362	

Grundlage: beobachtete Mittelwerte.
Der Fehlerterm ist Mittel der Quadrate(Fehler) = 1.007.
*. Die mittlere Differenz ist auf dem .05-Niveau signifikant.

10.3.18. Statistiken zum Kapitel 7.3.2

Tabelle 333: Reflexionskompetenztest – Residualer Lernzuwachs – Vergleich IG2 vs. KG

Gruppenstatistiken									
		gruppe	Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes		
SR Reflexionskompetenztest		Kein Feedback		21	-,3793577	,98844076	,21569546		
		Unterrichtsvideos		17	,1235048	1,12927214	,27388872		

Test bei unabhängigen Stichproben									
Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
SR Reflexionskompetenztest	Varianzen sind gleich	,336	,566	-1,463	36	,152	-,50286251	,34366424	Untere: -1,19984589 Obere: ,19412088
	Varianzen sind nicht gleich			-1,442	32,118	,159	-,50286251	,34862525	Untere: -1,21288691 Obere: ,20716189

Tabelle 334: Residualer Lernzuwachs – Test auf Normalverteilung – Ausschluss IG1

Kolmogorov-Smirnov-Anpassungstest									
		SR PCK	SR CK	SR Refl-Breite	SR Refl-Tiefe	SR Perspektive	SR Literatur	SR Refl-Gesamt	SR Refl-Test
N		38	38	38	38	37	37	38	38
Parameter der Normalverteilung ^{a,b}	Mittelwert	.0000000	.0000000	.0000000	.0000000	.0000000	.0000000	.0000000	.0000000
	Std.-Abweichung	.9863939	.9863939	.98639392	.98639392	.98601330	.98601330	.98639392	.98639392
Extremste Differenzen	Absolut	.148	.167	.090	.145	.123	.167	.159	.112
	Positiv	.085	.167	.065	.145	.123	.167	.159	.103
	Negativ	-.148	-.138	-.090	-.075	-.075	-.084	-.065	-.112
Statistik für Test		.148	.167	.090	.145	.123	.167	.159	.112
Asymptotische Signifikanz (2-seitig)		.035 ^c	.009 ^c	.200 ^{c,d}	.041 ^c	.172 ^c	.011 ^c	.016 ^c	.200 ^{c,d}

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

10. Anhang

c. Signifikanzkorrektur nach Lilliefors.

d. Dies ist eine untere Grenze der echten Signifikanz.

Tabelle 335: Residualer Lernzuwachs – Parametrische Korrelationen – Ausschluss IG1

		Korrelationen							
		SR PCK	SR CK	SR ReflBreite	SR Refl-Tiefe	SR Perspektiv	SR Literatur	SR Refl-Gesamt	SR Refl-Test
SR PCK	Korrelation nach Pearson	1	.253	-.002	.008	-.037	-.126	.012	.325*
	Signifikanz (2-seitig)		.125	.988	.960	.826	.457	.943	.047
	N	38	38	38	38	37	37	38	38
SR CK	Korrelation nach Pearson	.253	1	.045	-.088	.084	.040	-.086	.244
	Signifikanz (2-seitig)	.125		.791	.599	.621	.812	.607	.139
	N	38	38	38	38	37	37	38	38
SR ReflBreite	Korrelation nach Pearson	-.002	.045	1	.462**	.526**	.502**	.512**	.002
	Signifikanz (2-seitig)	.988	.791		.004	.001	.002	.001	.989
	N	38	38	38	38	37	37	38	38
SR ReflTiefe	Korrelation nach Pearson	.008	-.088	.462**	1	.674**	.740**	.997**	.040
	Signifikanz (2-seitig)	.960	.599	.004		.000	.000	.000	.812
	N	38	38	38	38	37	37	38	38
SR Perspektiv	Korrelation nach Pearson	-.037	.084	.526**	.674**	1	.922**	.676**	.074
	Signifikanz (2-seitig)	.826	.621	.001	.000		.000	.000	.662
	N	37	37	37	37	37	37	37	37
SR Literatur	Korrelation nach Pearson	-.126	.040	.502**	.740**	.922**	1	.737**	.113
	Signifikanz (2-seitig)	.457	.812	.002	.000	.000		.000	.505
	N	37	37	37	37	37	37	37	37
SR Refl-Gesamt	Korrelation nach Pearson	.012	-.086	.512**	.997**	.676**	.737**	1	.043
	Signifikanz (2-seitig)	.943	.607	.001	.000	.000	.000		.796
	N	38	38	38	38	37	37	38	38
SR ReflTest	Korrelation nach Pearson	.325*	.244	.002	.040	.074	.113	.043	1
	Signifikanz (2-seitig)	.047	.139	.989	.812	.662	.505	.796	
	N	38	38	38	38	37	37	38	38

*. Die Korrelation ist auf dem Niveau von 0.05 (2-seitig) signifikant.

**. Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant.

Tabelle 336: Residualer Lernzuwachs – Nichtparametrische Korrelationen – Ausschluss IG1

			Korrelationen							
			SR PCK	SR CK	SR ReflBreite	SR Refl-Tiefe	SR Perspektive	SR Literatur	SR Refl-Gesamt	SR Refl-Test
Spearman-Rho	SR PCK	Korrelationskoeffizient	1.000	.260	-.062	-.088	.032	-.091	-.058	.414**
		Sig. (2-seitig)	.	.115	.711	.601	.853	.591	.730	.010
		N	38	38	38	38	37	37	38	38
	SR CK	Korrelationskoeffizient	.260	1.000	.003	-.026	.091	.013	-.006	.297
		Sig. (2-seitig)	.115	.	.985	.876	.593	.938	.971	.070
		N	38	38	38	38	37	37	38	38
	SR ReflBreite	Korrelationskoeffizient	-.062	.003	1.000	.483**	.515**	.515**	.569**	.054
		Sig. (2-seitig)	.711	.985	.	.002	.001	.001	.000	.749
		N	38	38	38	38	37	37	38	38
	SR ReflTiefe	Korrelationskoeffizient	-.088	-.026	.483**	1.000	.648**	.701**	.986**	-.021
		Sig. (2-seitig)	.601	.876	.002	.	.000	.000	.000	.898
		N	38	38	38	38	37	37	38	38
	SR Perspektive	Korrelationskoeffizient	.032	.091	.515**	.648**	1.000	.939**	.656**	.027
		Sig. (2-seitig)	.853	.593	.001	.000	.	.000	.000	.873
		N	37	37	37	37	37	37	37	37
	SR Literatur	Korrelationskoeffizient	-.091	.013	.515**	.701**	.939**	1.000	.703**	.026
		Sig. (2-seitig)	.591	.938	.001	.000	.000	.	.000	.878
		N	37	37	37	37	37	37	37	37
	SR Refl-Gesamt	Korrelationskoeffizient	-.058	-.006	.569**	.986**	.656**	.703**	1.000	-.002
		Sig. (2-seitig)	.730	.971	.000	.000	.000	.000	.	.993

	N	38	38	38	38	37	37	38	38
SR ReflTest	Korrelationskoeffizient	.414**	.297	.054	-.021	.027	.026	-.002	1.000
	Sig. (2-seitig)	.010	.070	.749	.898	.873	.878	.993	.
	N	38	38	38	38	37	37	38	38

** Die Korrelation ist auf dem 0.01 Niveau signifikant (zweiseitig).

10.3.19. Statistiken zum Kapitel 7.3.3

Tabelle 337: Kodiermanual – Perspektivenübernahme – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest					xPerspektiven	yPerspektiven	zPerspektiven
N					57	57	58
Parameter der Normalverteilung ^{a,b}	Mittelwert				7.1930	7.2281	10.7931
	Std.-Abweichung				5.42625	4.29708	5.11537
Extremste Differenzen	Absolut				.160	.142	.165
	Positiv				.160	.142	.165
	Negativ				-.127	-.077	-.088
Statistik für Test					.160	.142	.165
Asymptotische Signifikanz (2-seitig)					.001 ^c	.006 ^c	.000 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

Tabelle 338: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Nichtparametrisch

Statistik für Test ^a			
	yPerspektiven - xPerspektiven	zPerspektiven - xPerspektiven	zPerspektiven - yPerspektiven
Z	-.150 ^b	-4.166 ^b	-5.176 ^b
Asymptotische Signifikanz (2-seitig)	.881	.000	.000

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 339: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Parametrisch

Test bei gepaarten Stichproben									
		Mittelwert	Std.-Abweichung	Gepaarte Differenzen		T	df	Sig. (2-seitig)	
				Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz				
					Untere	Obere			
Paaren 1	xPerspektiven - yPerspektiven	-.05357	5.40511	.72229	-1.50107	1.39393	-.074	55	.941
Paaren 2	xPerspektiven - zPerspektiven	-3.64912	5.97761	.79175	-5.23520	-2.06305	-4.609	56	.000
Paaren 3	yPerspektiven - zPerspektiven	-3.64912	4.23629	.56111	-4.77316	-2.52508	-6.503	56	.000

Tabelle 340: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Nichtparametrisch – Feedbackgruppe

Statistik für Test ^a			
	yPerspektiven - xPerspektiven	zPerspektiven - xPerspektiven	zPerspektiven - yPerspektiven
Z	-.334 ^b	-2.857 ^c	-3.227 ^c
Asymptotische Signifikanz (2-seitig)	.738	.004	.001

a. Wilcoxon-Test

b. Basiert auf positiven Rängen.

c. Basiert auf negativen Rängen.

Tabelle 341: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Parametrisch – Feedbackgruppe

Test bei gepaarten Stichproben								
		Mittelwert	Std.-Abweichung	Gepaarte Differenzen		T	df	Sig. (2-seitig)
				Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz			
					Untere	Obere		
Paaren 1	xPerspektiven - yPerspektiven	.35000	5.19387	1.16139	-2.08081	2.78081	.301	.766

10. Anhang

Paaren 2	xPerspektiven - zPerspektiven	-3.65000	5.47025	1.22319	-6.21016	-1.08984	-2.984	19	.008
Paaren 3	yPerspektiven - zPerspektiven	-4.00000	4.41290	.98675	-6.06530	-1.93470	-4.054	19	.001

Tabelle 342: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Nichtparametrisch – Selbstreflexionsgruppe

Z	Statistik für Test ^a		
	yPerspektiven - xPerspektiven	zPerspektiven - xPerspektiven	zPerspektiven - yPerspektiven
Z	-.687 ^b	-2.820 ^b	-3.622 ^b
Asymptotische Signifikanz (2-seitig)	.492	.005	.000

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 343: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Parametrisch – Feedbackgruppe

Test bei gepaarten Stichproben								
		Gepaarte Differenzen						
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	Sig. (2-seitig)
					Untere	Obere		
Paaren 1	xPerspektiven - yPerspektiven	-.73684	6.05385	1.38885	-3.65471	2.18102	-.531	.602
Paaren 2	xPerspektiven - zPerspektiven	-5.00000	6.71996	1.50263	-8.14504	-1.85496	-3.328	.004
Paaren 3	yPerspektiven - zPerspektiven	-4.45000	3.97988	.88993	-6.31264	-2.58736	-5.000	.000

Tabelle 344: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Nichtparametrisch – Kontrollgruppe

Z	Statistik für Test ^a		
	yPerspektiven - xPerspektiven	zPerspektiven - xPerspektiven	zPerspektiven - yPerspektiven
Z	-.399 ^b	-1.594 ^c	-1.770 ^c
Asymptotische Signifikanz (2-seitig)	.690	.111	.077

a. Wilcoxon-Test

b. Basiert auf positiven Rängen.

c. Basiert auf negativen Rängen.

Tabelle 345: Kodiermanual – Perspektivenübernahme – Wissenszuwachs – Parametrisch – Kontrollgruppe

Test bei gepaarten Stichproben								
		Gepaarte Differenzen						
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	Sig. (2-seitig)
					Untere	Obere		
Paaren 1	xPerspektiven - yPerspektiven	.23529	5.12993	1.24419	-2.40227	2.87286	.189	.852
Paaren 2	xPerspektiven - zPerspektiven	-2.05882	5.56182	1.34894	-4.91845	.80080	-1.526	.146
Paaren 3	yPerspektiven - zPerspektiven	-2.29412	4.23917	1.02815	-4.47370	-.11454	-2.231	.040

Residuen:

Tabelle 346: Perspektivenübernahme – Residualer Lernzuwachs – IG1 vs. IG2

Gruppenstatistiken				
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung
SR Perspektiven	Feedback	20	.0757496	.65938079
	Kein Feedback	20	.0764443	1,25205580

Test bei unabhängigen Stichproben

Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
SR Perspektiven	Varianzen sind gleich	7,681	,009	-,002	38	,998	-,00069471	,31641956	-,64125262	,63986320
	Varianzen sind nicht gleich			-,002	28,786	,998	-,00069471	,31641956	-,64805399	,64666456

Tabelle 347: Perspektivenübernahme – Residualer Lernzuwachs – IG1 vs. KG

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Perspektiven	Feedback	20	,0757496	,65938079	,14744203
	Unterrichtsvideos	17	-,1790517	1,00552364	,24387531

Test bei unabhängigen Stichproben										
Levene-Test der Varianzgleichheit				T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
SR Perspektiven	Varianzen sind gleich	2,225	,145	,924	35	,362	,25480139	,27565209	-,30480212	,81440489
	Varianzen sind nicht gleich			,894	26,817	,379	,25480139	,28498125	-,33011856	,83972133

Literaturbezug:

Tabelle 348: Kodiermanual – Literaturbezug – Test auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest				
		xPersLiteratur	yPersLiteratur	zPersLiteratur
N		57	57	58
Parameter der Normalverteilung ^{a,b}	Mittelwert	7.4386	7.7193	13.5345
	Std.-Abweichung	5.73840	4.92354	6.94645
Extremste Differenzen	Absolut	.163	.143	.142
	Positiv	.163	.143	.142
	Negativ	-.131	-.088	-.075
Statistik für Test		.163	.143	.142
Asymptotische Signifikanz (2-seitig)		.001 ^c	.005 ^c	.005 ^c

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

c. Signifikanzkorrektur nach Lilliefors.

Tabelle 349: Kodiermanual – Literaturbezug – Wissenszuwachs – Nichtparametrisch

Statistik für Test ^a			
	yPersLiteratur - xPersLiteratur	zPersLiteratur - xPersLiteratur	zPersLiteratur - yPersLiteratur
Z	-.376 ^b	-4.970 ^b	-5.751 ^b
Asymptotische Signifikanz (2-seitig)	.707	.000	.000

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Tabelle 350: Kodiermanual – Literaturbezug – Wissenszuwachs – Parametrisch

Test bei gepaarten Stichproben									
			Gepaarte Differenzen		95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	xPersLiteratur - yPersLiteratur	-.30357	5.65016	.75503	-1.81670	1.20955	-.402	55	.689
Paaren 2	xPersLiteratur - zPersLiteratur	-6.14035	7.75159	1.02672	-8.19712	-4.08358	-5.981	56	.000

10. Anhang

Paaren 3	yPersLiteratur - zPersLiteratur	-5.89474	5.92418	.78468	-7.46663	-4.32284	-7.512	56	.000

Tabelle 351: Kodiermanual – Literaturbezug – Wissenszuwachs – Nichtparametrisch – Feedbackgruppe

Statistik für Test ^a			
	yPersLiteratur - xPersLiteratur	zPersLiteratur - xPersLiteratur	zPersLiteratur - yPersLiteratur
Z	-.570 ^b	-3.144 ^c	-3.778 ^c
Asymptotische Signifikanz (2-seitig)	.569	.002	.000

a. Wilcoxon-Test
b. Basiert auf positiven Rängen.
c. Basiert auf negativen Rängen.

Tabelle 352: Kodiermanual – Literaturbezug – Wissenszuwachs – Nichtparametrisch – Selbstreflexionsgruppe

Statistik für Test ^a			
	yPersLiteratur - xPersLiteratur	zPersLiteratur - xPersLiteratur	zPersLiteratur - yPersLiteratur
Z	-.677 ^b	-3.163 ^b	-3.704 ^b
Asymptotische Signifikanz (2-seitig)	.499	.002	.000

a. Wilcoxon-Test
b. Basiert auf negativen Rängen.

Tabelle 353: Kodiermanual – Literaturbezug – Wissenszuwachs – Nichtparametrisch – Kontrollgruppe

Statistik für Test ^a			
	yPersLiteratur - xPersLiteratur	zPersLiteratur - xPersLiteratur	zPersLiteratur - yPersLiteratur
Z	-.171 ^b	-2.276 ^b	-2.232 ^b
Asymptotische Signifikanz (2-seitig)	.864	.023	.026

a. Wilcoxon-Test
b. Basiert auf negativen Rängen.

Tabelle 354: Kodiermanual – Literaturbezug – Wissenszuwachs – Parametrisch – Feedbackgruppe

Test bei gepaarten Stichproben								
		Gepaarte Differenzen						
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz			
					Untere	Obere	T	Sig. (2-seitig)
Paaren 1	xPersLiteratur - yPersLiteratur	.60000	5.42315	1.21265	-1.93811	3.13811	.495	.626
Paaren 2	xPersLiteratur - zPersLiteratur	-5.75000	7.10726	1.58923	-9.07630	-2.42370	-3.618	.002
Paaren 3	yPersLiteratur - zPersLiteratur	-6.35000	5.64078	1.26132	-8.98997	-3.71003	-5.034	.000

Tabelle 355: Kodiermanual – Literaturbezug – Wissenszuwachs – Parametrisch – Selbstreflexionsgruppe

Test bei gepaarten Stichproben								
		Gepaarte Differenzen						
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz			
					Untere	Obere	T	Sig. (2-seitig)
Paaren 1	xPersLiteratur - yPersLiteratur	-1.10526	6.52257	1.49638	-4.24904	2.03851	-.739	.470
Paaren 2	xPersLiteratur - zPersLiteratur	-8.05000	8.54385	1.91046	-12.04864	-4.05136	-4.214	.000
Paaren 3	yPersLiteratur - zPersLiteratur	-7.15000	6.01117	1.34414	-9.96332	-4.33668	-5.319	.000

Tabelle 356: Kodiermanual – Literaturbezug – Wissenszuwachs – Parametrisch – Kontrollgruppe

Test bei gepaarten Stichproben									
		Gepaarte Differenzen							
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	xPersLiteratur - yPersLiteratur	-.47059	5.01395	1.21606	-3.04852	2.10735	-.387	16	.704
Paaren 2	xPersLiteratur - zPersLiteratur	-4.35294	7.44934	1.80673	-8.18304	-.52284	-2.409	16	.028
Paaren 3	yPersLiteratur - zPersLiteratur	-3.88235	5.96744	1.44732	-6.95053	-.81418	-2.682	16	.016

Tabelle 357: Literaturbezug – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. IG2

Statistik für Test ^a	
	SR Literatur
Mann-Whitney-U	194,500
Wilcoxon-W	404,500
Z	-.149
Asymptotische Signifikanz (2-seitig)	,882
Exakte Signifikanz [2*(1-seitige Sig.)]	,883 ^b

a. Gruppenvariable: gruppe Gruppenzuordnung
b. Nicht für Bindungen korrigiert.

Tabelle 358: Literaturbezug – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG

Statistik für Test ^a	
	SR Literatur
Mann-Whitney-U	131,500
Wilcoxon-W	284,500
Z	-1,174
Asymptotische Signifikanz (2-seitig)	,241
Exakte Signifikanz [2*(1-seitige Sig.)]	,244 ^b

a. Gruppenvariable: gruppe Gruppenzuordnung
b. Nicht für Bindungen korrigiert.

Tabelle 359: Literaturbezug – Residualer Lernzuwachs – Parametrisch – IG1 vs. IG2

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Literatur	Feedback	20	,0125249	,69996468	,15651686
	Kein Feedback	20	,1296548	1,18205601	,26431576

Test bei unabhängigen Stichproben									
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
SR Literatur	Varianzen sind gleich	8,872	,005	-,381	38	,705	-,11712989	,30718129	-,73898591
	Varianzen sind nicht gleich			-,381	30,866	,706	-,11712989	,30718129	-,74374070

Tabelle 360: Literaturbezug – Residualer Lernzuwachs – Nichtparametrisch – IG1 vs. KG

Gruppenstatistiken					
	gruppe Gruppenzuordnung	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SR Literatur	Feedback	20	,0125249	,69996468	,15651686
	Unterrichtsvideos	17	-,1672702	1,07028251	,25958164

10. Anhang

Test bei unabhängigen Stichproben									
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz
									Untere Obere
SR Literatur	Varianzen sind gleich	2,749	,106	,613	35	,544	,17979514	,29313919	-,41530906 ,77489934
	Varianzen sind nicht gleich			,593	26,769	,558	,17979514	,30311739	-,44240146 ,80199174

10.3.20. Statistiken zum Kapitel 7.3.4

Tabelle 361: Extremgruppenvergleich – Einfluss PCK-Test auf Reflexionskompetenztest

Tests der Zwischensubjekteffekte						
Abhängige Variable: Average Percentage Refl Post						
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	860,251 ^a	2	430,126	1,035	,362	,036
Konstanter Term	17827,810	1	17827,810	42,882	,000	,438
PerReflPre	860,251	1	860,251	2,069	,156	,036
DichoPCK	1,881	1	1,881	,005	,947	,000
Fehler	22865,802	55	415,742			
Gesamt	243055,556	58				
Korrigierte Gesamtvariation	23726,054	57				

a. R-Quadrat = ,036 (korrigiertes R-Quadrat = ,001)

Tabelle 362: Extremgruppenvergleich – Einfluss PCK-Test auf Gesamtpunktzahl des Kodiermanuals

Tests der Zwischensubjekteffekte						
Abhängige Variable: zGesamtpunktzahl						
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	39196,731 ^a	2	19598,365	29,947	,000	,521
Konstanter Term	4116,053	1	4116,053	6,289	,015	,103
xGesamtpunktzahl	39150,110	1	39150,110	59,822	,000	,521
DichoPCK	307,513	1	307,513	,470	,496	,008
Fehler	35994,166	55	654,439			
Gesamt	411416,000	58				
Korrigierte Gesamtvariation	75190,897	57				

a. R-Quadrat = ,521 (korrigiertes R-Quadrat = ,504)

Tabelle 363: Extremgruppenvergleich – Einfluss CK-Test auf Reflexionskompetenztest

Tests der Zwischensubjekteffekte						
Abhängige Variable: Average Percentage Refl Post						
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	1299,451 ^a	2	649,726	1,593	,212	,055
Konstanter Term	18267,794	1	18267,794	44,801	,000	,449
PerReflPre	719,950	1	719,950	1,766	,189	,031
DichoCK	441,081	1	441,081	1,082	,303	,019
Fehler	22426,602	55	407,756			
Gesamt	243055,556	58				
Korrigierte Gesamtvariation	23726,054	57				

a. R-Quadrat = ,055 (korrigiertes R-Quadrat = ,020)

Tabelle 364: Extremgruppenvergleich – Einfluss PCK-Gesamtpunktzahl Kodiermanual

Tests der Zwischensubjekteffekte

Abhängige Variable: zGesamtpunktzahl

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	38960,293 ^a	2	19480,146	29,572	,000	,518
Konstanter Term	4394,407	1	4394,407	6,671	,012	,108
xGesamtpunktzahl	38951,948	1	38951,948	59,131	,000	,518
DichoCK	71,075	1	71,075	,108	,744	,002
Fehler	36230,604	55	658,738			
Gesamt	411416,000	58				
Korrigierte Gesamtvariation	75190,897	57				

a. R-Quadrat = ,518 (korrigiertes R-Quadrat = ,501)

Tabelle 365: Einfluss des Vortestergebnisses des PCK-Tests auf die Leistungsentwicklung im Reflexionskompetenztest

Tests der Zwischensubjekteffekte

Abhängige Variable: Average Percentage Refl Post

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	867,236 ^a	2	433,618	1,043	,359	,037
Konstanter Term	3726,680	1	3726,680	8,967	,004	,140
PerReflPre	785,115	1	785,115	1,889	,175	,033
xCores	8,865	1	8,865	,021	,884	,000
Fehler	22858,818	55	415,615			
Gesamt	243055,556	58				
Korrigierte Gesamtvariation	23726,054	57				

a. R-Quadrat = ,037 (korrigiertes R-Quadrat = ,002)

Tabelle 366: Einfluss des Vortestergebnisses des PCK-Tests auf die Entwicklung der Gesamtpunktzahl des Kodiermanuals

Tests der Zwischensubjekteffekte

Abhängige Variable: zGesamtpunktzahl

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	38889,725 ^a	2	19444,863	29,461	,000	,517
Konstanter Term	800,998	1	800,998	1,214	,275	,022
xGesamtpunktzahl	36220,616	1	36220,616	54,878	,000	,499
xCores	,508	1	,508	,001	,978	,000
Fehler	36301,171	55	660,021			
Gesamt	411416,000	58				
Korrigierte Gesamtvariation	75190,897	57				

a. R-Quadrat = ,517 (korrigiertes R-Quadrat = ,500)

Tabelle 367: Einfluss des Vortestergebnisses des CK-Tests auf die Leistungsentwicklung im Reflexionskompetenztest

Tests der Zwischensubjekteffekte

Abhängige Variable: Average Percentage Refl Post

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	866,263 ^a	2	433,131	1,042	,360	,037
Konstanter Term	9847,524	1	9847,524	23,693	,000	,301
PerReflPre	865,626	1	865,626	2,083	,155	,036
PerCKPre	7,892	1	7,892	,019	,891	,000
Fehler	22859,791	55	415,633			
Gesamt	243055,556	58				
Korrigierte Gesamtvariation	23726,054	57				

a. R-Quadrat = ,037 (korrigiertes R-Quadrat = ,001)

10. Anhang

Tabelle 368: Einfluss des Vortestergebnisses des CK-Tests auf die Entwicklung der Gesamtpunktzahl des Kodiermanuals

Tests der Zwischensubjekteffekte						
Abhängige Variable: zGesamtpunktzahl						
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	39556,700 ^a	2	19778,350	30,527	,000	,526
Konstanter Term	435,098	1	435,098	,672	,416	,012
PerCKPre	667,483	1	667,483	1,030	,315	,018
xGesamtpunktzahl	39225,576	1	39225,576	60,543	,000	,524
Fehler	35634,196	55	647,894			
Gesamt	411416,000	58				
Korrigierte Gesamtvariation	75190,897	57				

a. R-Quadrat = ,526 (korrigiertes R-Quadrat = ,509)

Danksagung

Am Ende möchte ich mich bei den Menschen in meinem Leben bedanken, ohne die dieser Lebensabschnitt nicht möglich gewesen wäre. Zunächst ein herzlicher Dank an meine Prüfer Prof. Dr. Astrid Rank, Prof. Dr. Oliver Tepner, Prof. Dr. Stefan Krauss und Prof. Dr. Arne Dittmer. Vielen Dank für Ihre Bereitschaft, sich mit meinem Thema zu befassen und mich zu prüfen, aber auch für die zahlreichen Möglichkeiten, Ihnen meine Arbeit während der Entstehung vorzustellen und für die damit verbundenen sehr guten Ratschläge. Doch ganz speziell möchte ich mich bei dir, lieber Oli bedanken. Deine Unterstützung für mich bei dieser Arbeit hätte nicht besser sein können. Am meisten möchte ich mich jedoch für dein Verständnis für mich und alle meine Lebenslagen bedanken.

Weiter möchte ich mich bei der besten Arbeitsgruppe der Welt bedanken. Victoria, Nadine, Elmi, Lars und Flo, ihr wart über die ganze Zeit meine Anlaufstelle für diverse Fragen. Wie schön ist es, in einer Arbeitsgruppe arbeiten zu dürfen, in der jeder dem Anderen zu jeder Zeit hilft. Nicht zu vergessen sind hier auch meine ehemaligen Büromitglieder Stefan und Martina und die gute Seele der Chemiedidaktik, Sandra. Aber auch ein Dank an alle SHKs, im speziellen ein Dank an Linda und Richie. Wenn ich den Titel teilen könnte, würde ich euch etwas davon abgeben. Ihr habt mir stundenlang geholfen, mich beraten und mich aufgebaut.

Ebenso geht der Dank an alle meine Studierenden. Besonders möchte ich hierbei Theresa erwähnen, deren Zulassungsarbeit Grundlage für mein Kodiermanual wurde. Danke an alle Schülerinnen und Schüler, welche an den Schülertagen so perfekt mitgewirkt haben, aber auch an meine Schulleiter Herr Langer und Frau Retzer, dass ich meine Arbeit während meiner Lehrtätigkeit an ihrer Schule fertigstellen durfte.

Ein Dank auch meinen Freunden Judith, Maria, Nina, Steffi H., Steffi A., Vera, Susi B., Anja, Lena, Jürgen, Kirsten, Martin und Sue. Ihr habt immer Interesse gezeigt und mich durch eure Nachfragen unterstützt und motiviert. Speziell möchte ich Anna danken. Du hast tatsächlich alles durchgelesen, allein das ist schon eine Leistung, aber danke zusätzlich für die vielen tollen Verbesserungsvorschläge.

Auch meiner Familie möchte ich für ihr Verständnis danken. Danke Mama, für die vielen Male wo du meinen Hansi beaufsichtigt hast. Danke Mia, auch du musstest durch das ganze Buch und hast mir geholfen, alle Fehler zu verbessern. Die Lücke die hier in unserer Familie aufging konntest du schließen. Auch danke an dich, Anja. Mit dir konnte ich schreiben, telefonieren und reden. Du hast mir zugehört und mich unterstützt. Deine Hilfe über die ganze Zeit ist für mich von unschätzbarem Wert.

Am Ende möchte ich den drei stärksten Männern in meinem Leben danken. Dieses Buch ist für dich Papa, du hast immer an mich geglaubt und alle Selbstzweifel beseitigt. Du fehlst mir jeden einzelnen Tag! Danke, kleiner Hansi, deine fröhliche, liebe Art lassen alle Probleme sofort vergessen. Und danke, meinem Mann Oli. Du hast nicht nur meine Studie diverse Male mit deinen Klassen unterstützt, deine liebevolle, verständnisvolle Art und der Glaube an mich, deine stetige Bereitschaft unser Kind zu beschäftigen, so dass ich schreiben konnte, bedeutet mir alles.

Bisher erschienene Bände der Reihe „*Studien zum Physik- und Chemielernen*“

ISSN 1614-8967 (vormals *Studien zum Physiklernen* ISSN 1435-5280)

- 1 Helmut Fischler, Jochen Peuckert (Hrsg.): Concept Mapping in fachdidaktischen Forschungsprojekten der Physik und Chemie
ISBN 978-3-89722-256-4 40.50 EUR
- 2 Anja Schoster: Bedeutungsentwicklungsprozesse beim Lösen algorithmischer Physikaufgaben. *Eine Fallstudie zu Lernprozessen von Schülern im Physiknachhilfeunterricht während der Bearbeitung algorithmischer Physikaufgaben*
ISBN 978-3-89722-045-4 40.50 EUR
- 3 Claudia von Aufschnaiter: Bedeutungsentwicklungen, Interaktionen und situatives Erleben beim Bearbeiten physikalischer Aufgaben
ISBN 978-3-89722-143-7 40.50 EUR
- 4 Susanne Haerberlen: Lernprozesse im Unterricht mit Wasserstromkreisen. *Eine Fallstudie in der Sekundarstufe I*
ISBN 978-3-89722-172-7 40.50 EUR
- 5 Kerstin Haller: Über den Zusammenhang von Handlungen und Zielen. *Eine empirische Untersuchung zu Lernprozessen im physikalischen Praktikum*
ISBN 978-3-89722-242-7 40.50 EUR
- 6 Michaela Horstendahl: Motivationale Orientierungen im Physikunterricht
ISBN 978-3-89722-227-4 50.00 EUR
- 7 Stefan Deylitz: Lernergebnisse in der Quanten-Atomphysik. *Evaluation des Bremer Unterrichtskonzepts*
ISBN 978-3-89722-291-5 40.50 EUR
- 8 Lorenz Hucke: Handlungsregulation und Wissenserwerb in traditionellen und computergestützten Experimenten des physikalischen Praktikums
ISBN 978-3-89722-316-5 50.00 EUR
- 9 Heike Theyßen: Ein Physikpraktikum für Studierende der Medizin. *Darstellung der Entwicklung und Evaluation eines adressatenspezifischen Praktikums nach dem Modell der Didaktischen Rekonstruktion*
ISBN 978-3-89722-334-9 40.50 EUR
- 10 Annette Schick: Der Einfluß von Interesse und anderen selbstbezogenen Kognitionen auf Handlungen im Physikunterricht. *Fallstudien zu Interessenhandlungen im Physikunterricht*
ISBN 978-3-89722-380-6 40.50 EUR
- 11 Roland Berger: Moderne bildgebende Verfahren der medizinischen Diagnostik. *Ein Weg zu interessanterem Physikunterricht*
ISBN 978-3-89722-445-2 40.50 EUR

- 12 Johannes Werner: Vom Licht zum Atom. *Ein Unterrichtskonzept zur Quantenphysik unter Nutzung des Zeigermodells*
ISBN 978-3-89722-471-1 40.50 EUR
- 13 Florian Sander: Verbindung von Theorie und Experiment im physikalischen Praktikum. *Eine empirische Untersuchung zum handlungsbezogenen Vorverständnis und dem Einsatz grafikorientierter Modellbildung im Praktikum*
ISBN 978-3-89722-482-7 40.50 EUR
- 14 Jörn Gerdes: Der Begriff der physikalischen Kompetenz. *Zur Validierung eines Konstruktes*
ISBN 978-3-89722-510-7 40.50 EUR
- 15 Malte Meyer-Arndt: Interaktionen im Physikpraktikum zwischen Studierenden und Betreuern. *Feldstudie zu Bedeutungsentwicklungsprozessen im physikalischen Praktikum*
ISBN 978-3-89722-541-1 40.50 EUR
- 16 Dietmar Höttecke: Die Natur der Naturwissenschaften historisch verstehen. *Fachdidaktische und wissenschaftshistorische Untersuchungen*
ISBN 978-3-89722-607-4 40.50 EUR
- 17 Gil Gabriel Mavanga: Entwicklung und Evaluation eines experimentell- und phänomenorientierten Optikcurriculums. *Untersuchung zu Schülervorstellungen in der Sekundarstufe I in Mosambik und Deutschland*
ISBN 978-3-89722-721-7 40.50 EUR
- 18 Meike Ute Zastrow: Interaktive Experimentieranleitungen. *Entwicklung und Evaluation eines Konzeptes zur Vorbereitung auf das Experimentieren mit Messgeräten im Physikalischen Praktikum*
ISBN 978-3-89722-802-3 40.50 EUR
- 19 Gunnar Friege: Wissen und Problemlösen. *Eine empirische Untersuchung des wissenszentrierten Problemlösens im Gebiet der Elektrizitätslehre auf der Grundlage des Experten-Novizen-Vergleichs*
ISBN 978-3-89722-809-2 40.50 EUR
- 20 Erich Starauschek: Physikunterricht nach dem Karlsruher Physikkurs. *Ergebnisse einer Evaluationsstudie*
ISBN 978-3-89722-823-8 40.50 EUR
- 21 Roland Paatz: Charakteristika analogiebasierten Denkens. *Vergleich von Lernprozessen in Basis- und Zielbereich*
ISBN 978-3-89722-944-0 40.50 EUR
- 22 Silke Mikelskis-Seifert: Die Entwicklung von Metakzepten zur Teilchenvorstellung bei Schülern. *Untersuchung eines Unterrichts über Modelle mithilfe eines Systems multipler Repräsentationsebenen*
ISBN 978-3-8325-0013-9 40.50 EUR
- 23 Brunhild Landwehr: Distanzen von Lehrkräften und Studierenden des Sachunterrichts zur Physik. *Eine qualitativ-empirische Studie zu den Ursachen*
ISBN 978-3-8325-0044-3 40.50 EUR

- 24 Lydia Murmann: Physiklernen zu Licht, Schatten und Sehen. *Eine phänomenografische Untersuchung in der Primarstufe*
ISBN 978-3-8325-0060-3 40.50 EUR
- 25 Thorsten Bell: Strukturprinzipien der Selbstregulation. *Komplexe Systeme, Elementarisierungen und Lernprozessstudien für den Unterricht der Sekundarstufe II*
ISBN 978-3-8325-0134-1 40.50 EUR
- 26 Rainer Müller: Quantenphysik in der Schule
ISBN 978-3-8325-0186-0 40.50 EUR
- 27 Jutta Roth: Bedeutungsentwicklungsprozesse von Physikerinnen und Physikern in den Dimensionen Komplexität, Zeit und Inhalt
ISBN 978-3-8325-0183-9 40.50 EUR
- 28 Andreas Saniter: Spezifika der Verhaltensmuster fortgeschrittener Studierender der Physik
ISBN 978-3-8325-0292-8 40.50 EUR
- 29 Thomas Weber: Kumulatives Lernen im Physikunterricht. *Eine vergleichende Untersuchung in Unterrichtsgängen zur geometrischen Optik*
ISBN 978-3-8325-0316-1 40.50 EUR
- 30 Markus Rehm: Über die Chancen und Grenzen moralischer Erziehung im naturwissenschaftlichen Unterricht
ISBN 978-3-8325-0368-0 40.50 EUR
- 31 Marion Budde: Lernwirkungen in der Quanten-Atom-Physik. *Fallstudien über Resonanzen zwischen Lernangeboten und SchülerInnen-Vorstellungen*
ISBN 978-3-8325-0483-0 40.50 EUR
- 32 Thomas Reyer: Oberflächenmerkmale und Tiefenstrukturen im Unterricht. *Exemplarische Analysen im Physikunterricht der gymnasialen Sekundarstufe*
ISBN 978-3-8325-0488-5 40.50 EUR
- 33 Christoph Thomas Müller: Subjektive Theorien und handlungsleitende Kognitionen von Lehrern als Determinanten schulischer Lehr-Lern-Prozesse im Physikunterricht
ISBN 978-3-8325-0543-1 40.50 EUR
- 34 Gabriela Jonas-Ahrend: Physiklehrervorstellungen zum Experiment im Physikunterricht
ISBN 978-3-8325-0576-9 40.50 EUR
- 35 Dimitrios Stavrou: Das Zusammenspiel von Zufall und Gesetzmäßigkeiten in der nicht-linearen Dynamik. *Didaktische Analyse und Lernprozesse*
ISBN 978-3-8325-0609-4 40.50 EUR
- 36 Katrin Engeln: Schülerlabors: authentische, aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken
ISBN 978-3-8325-0689-6 40.50 EUR
- 37 Susann Hartmann: Erklärungsvielfalt
ISBN 978-3-8325-0730-5 40.50 EUR

- 38 Knut Neumann: Didaktische Rekonstruktion eines physikalischen Praktikums für Physiker
ISBN 978-3-8325-0762-6 40.50 EUR
- 39 Michael Späth: Kontextbedingungen für Physikunterricht an der Hauptschule. *Möglichkeiten und Ansatzpunkte für einen fachübergreifenden, handlungsorientierten und berufsorientierten Unterricht*
ISBN 978-3-8325-0827-2 40.50 EUR
- 40 Jörg Hirsch: Interesse, Handlungen und situatives Erleben von Schülerinnen und Schülern beim Bearbeiten physikalischer Aufgaben
ISBN 978-3-8325-0875-3 40.50 EUR
- 41 Monika Hüther: Evaluation einer hypermedialen Lernumgebung zum Thema Gasgesetze. *Eine Studie im Rahmen des Physikpraktikums für Studierende der Medizin*
ISBN 978-3-8325-0911-8 40.50 EUR
- 42 Maike Tesch: Das Experiment im Physikunterricht. *Didaktische Konzepte und Ergebnisse einer Videostudie*
ISBN 978-3-8325-0975-0 40.50 EUR
- 43 Nina Nicolai: Skriptgeleitete Eltern-Kind-Interaktion bei Chemiehausaufgaben. *Eine Evaluationsstudie im Themenbereich Säure-Base*
ISBN 978-3-8325-1013-8 40.50 EUR
- 44 Antje Leisner: Entwicklung von Modellkompetenz im Physikunterricht
ISBN 978-3-8325-1020-6 40.50 EUR
- 45 Stefan Rumann: Evaluation einer Interventionsstudie zur Säure-Base-Thematik
ISBN 978-3-8325-1027-5 40.50 EUR
- 46 Thomas Wilhelm: Konzeption und Evaluation eines Kinematik/Dynamik-Lehrgangs zur Veränderung von Schülervorstellungen mit Hilfe dynamisch ikonischer Repräsentationen und graphischer Modellbildung – mit CD-ROM
ISBN 978-3-8325-1046-6 45.50 EUR
- 47 Andrea Maier-Richter: Computerunterstütztes Lernen mit Lösungsbeispielen in der Chemie. *Eine Evaluationsstudie im Themenbereich Löslichkeit*
ISBN 978-3-8325-1046-6 40.50 EUR
- 48 Jochen Peuckert: Stabilität und Ausprägung kognitiver Strukturen zum Atombegriff
ISBN 978-3-8325-1104-3 40.50 EUR
- 49 Maik Walpuski: Optimierung von experimenteller Kleingruppenarbeit durch Strukturierungshilfen und Feedback
ISBN 978-3-8325-1184-5 40.50 EUR
- 50 Helmut Fischler, Christiane S. Reiners (Hrsg.): Die Teilchenstruktur der Materie im Physik- und Chemieunterricht
ISBN 978-3-8325-1225-5 34.90 EUR
- 51 Claudia Eysel: Interdisziplinäres Lehren und Lernen in der Lehrerbildung. *Eine empirische Studie zum Kompetenzerwerb in einer komplexen Lernumgebung*
ISBN 978-3-8325-1238-5 40.50 EUR

- 52 Johannes Günther: Lehrerfortbildung über die Natur der Naturwissenschaften. *Studien über das Wissenschaftsverständnis von Grundschullehrkräften*
ISBN 978-3-8325-1287-3 40.50 EUR
- 53 Christoph Neugebauer: Lernen mit Simulationen und der Einfluss auf das Problemlösen in der Physik
ISBN 978-3-8325-1300-9 40.50 EUR
- 54 Andreas Schnirch: Gendergerechte Interessen- und Motivationsförderung im Kontext naturwissenschaftlicher Grundbildung. *Konzeption, Entwicklung und Evaluation einer multimedial unterstützten Lernumgebung*
ISBN 978-3-8325-1334-4 40.50 EUR
- 55 Hilde Köster: Freies Explorieren und Experimentieren. *Eine Untersuchung zur selbstbestimmten Gewinnung von Erfahrungen mit physikalischen Phänomenen im Sachunterricht*
ISBN 978-3-8325-1348-1 40.50 EUR
- 56 Eva Heran-Dörr: Entwicklung und Evaluation einer Lehrerfortbildung zur Förderung der physikdidaktischen Kompetenz von Sachunterrichtslehrkräften
ISBN 978-3-8325-1377-1 40.50 EUR
- 57 Agnes Szabone Varnai: Unterstützung des Problemlösens in Physik durch den Einsatz von Simulationen und die Vorgabe eines strukturierten Kooperationsformats
ISBN 978-3-8325-1403-7 40.50 EUR
- 58 Johannes Rethfeld: Aufgabenbasierte Lernprozesse in selbstorganisationsoffenem Unterricht der Sekundarstufe I zum Themengebiet ELEKTROSTATIK. *Eine Feldstudie in vier 10. Klassen zu einer kartenbasierten Lernumgebung mit Aufgaben aus der Elektrostatik*
ISBN 978-3-8325-1416-7 40.50 EUR
- 59 Christian Henke: Experimentell-naturwissenschaftliche Arbeitsweisen in der Oberstufe. *Untersuchung am Beispiel des HIGHSEA-Projekts in Bremerhaven*
ISBN 978-3-8325-1515-7 40.50 EUR
- 60 Lutz Kasper: Diskursiv-narrative Elemente für den Physikunterricht. *Entwicklung und Evaluation einer multimedialen Lernumgebung zum Erdmagnetismus*
ISBN 978-3-8325-1537-9 40.50 EUR
- 61 Thorid Rabe: Textgestaltung und Aufforderung zu Selbsterklärungen beim Physiklernen mit Multimedia
ISBN 978-3-8325-1539-3 40.50 EUR
- 62 Ina Glemnitz: Vertikale Vernetzung im Chemieunterricht. *Ein Vergleich von traditionellem Unterricht mit Unterricht nach Chemie im Kontext*
ISBN 978-3-8325-1628-4 40.50 EUR
- 63 Erik Einhaus: Schülerkompetenzen im Bereich Wärmelehre. *Entwicklung eines Testinstruments zur Überprüfung und Weiterentwicklung eines normativen Modells fachbezogener Kompetenzen*
ISBN 978-3-8325-1630-7 40.50 EUR

- 64 Jasmin Neuroth: Concept Mapping als Lernstrategie. *Eine Interventionsstudie zum Chemielernen aus Texten*
ISBN 978-3-8325-1659-8 40.50 EUR
- 65 Hans Gerd Hegeler-Burkhart: Zur Kommunikation von Hauptschülerinnen und Hauptschülern in einem handlungsorientierten und fächerübergreifenden Unterricht mit physikalischen und technischen Inhalten
ISBN 978-3-8325-1667-3 40.50 EUR
- 66 Karsten Rincke: Sprachentwicklung und Fachlernen im Mechanikunterricht. *Sprache und Kommunikation bei der Einführung in den Kraftbegriff*
ISBN 978-3-8325-1699-4 40.50 EUR
- 67 Nina Strehle: Das Ion im Chemieunterricht. *Alternative Schülervorstellungen und curriculare Konsequenzen*
ISBN 978-3-8325-1710-6 40.50 EUR
- 68 Martin Hopf: Problemorientierte Schülerexperimente
ISBN 978-3-8325-1711-3 40.50 EUR
- 69 Anne Beerenwinkel: Fostering conceptual change in chemistry classes using expository texts
ISBN 978-3-8325-1721-2 40.50 EUR
- 70 Roland Berger: Das Gruppenpuzzle im Physikunterricht der Sekundarstufe II. *Eine empirische Untersuchung auf der Grundlage der Selbstbestimmungstheorie der Motivation*
ISBN 978-3-8325-1732-8 40.50 EUR
- 71 Giuseppe Colicchia: Physikunterricht im Kontext von Medizin und Biologie. *Entwicklung und Erprobung von Unterrichtseinheiten*
ISBN 978-3-8325-1746-5 40.50 EUR
- 72 Sandra Winheller: Geschlechtsspezifische Auswirkungen der Lehrer-Schüler-Interaktion im Chemieanfangsunterricht
ISBN 978-3-8325-1757-1 40.50 EUR
- 73 Isabel Wahser: Training von naturwissenschaftlichen Arbeitsweisen zur Unterstützung experimenteller Kleingruppenarbeit im Fach Chemie
ISBN 978-3-8325-1815-8 40.50 EUR
- 74 Claus Brell: Lernmedien und Lernerfolg - reale und virtuelle Materialien im Physikunterricht. *Empirische Untersuchungen in achten Klassen an Gymnasien (Laborstudie) zum Computereinsatz mit Simulation und IBE*
ISBN 978-3-8325-1829-5 40.50 EUR
- 75 Rainer Wackermann: Überprüfung der Wirksamkeit eines Basismodell-Trainings für Physiklehrer
ISBN 978-3-8325-1882-0 40.50 EUR
- 76 Oliver Tepner: Effektivität von Aufgaben im Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-1919-3 40.50 EUR

- 77 Claudia Geyer: Museums- und Science-Center-Besuche im naturwissenschaftlichen Unterricht aus einer motivationalen Perspektive. *Die Sicht von Lehrkräften und Schülerinnen und Schülern*
ISBN 978-3-8325-1922-3 40.50 EUR
- 78 Tobias Leonhard: Professionalisierung in der Lehrerbildung. *Eine explorative Studie zur Entwicklung professioneller Kompetenzen in der Lehrererstausbildung*
ISBN 978-3-8325-1924-7 40.50 EUR
- 79 Alexander Kauertz: Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben
ISBN 978-3-8325-1925-4 40.50 EUR
- 80 Regina Hübinger: Schüler auf Weltreise. *Entwicklung und Evaluation von Lehr-/Lernmaterialien zur Förderung experimentell-naturwissenschaftlicher Kompetenzen für die Jahrgangsstufen 5 und 6*
ISBN 978-3-8325-1932-2 40.50 EUR
- 81 Christine Waltner: Physik lernen im Deutschen Museum
ISBN 978-3-8325-1933-9 40.50 EUR
- 82 Torsten Fischer: Handlungsmuster von Physiklehrkräften beim Einsatz neuer Medien. *Fallstudien zur Unterrichtspraxis*
ISBN 978-3-8325-1948-3 42.00 EUR
- 83 Corinna Kieren: Chemiehausaufgaben in der Sekundarstufe I des Gymnasiums. *Fragebogenerhebung zur gegenwärtigen Praxis und Entwicklung eines optimierten Hausaufgabendesigns im Themenbereich Säure-Base*
978-3-8325-1975-9 37.00 EUR
- 84 Marco Thiele: Modelle der Thermohalinen Zirkulation im Unterricht. *Eine empirische Studie zur Förderung des Modellverständnisses*
ISBN 978-3-8325-1982-7 40.50 EUR
- 85 Bernd Zinn: Physik lernen, um Physik zu lehren. *Eine Möglichkeit für interessanteren Physikunterricht*
ISBN 978-3-8325-1995-7 39.50 EUR
- 86 Esther Klaes: Außerschulische Lernorte im naturwissenschaftlichen Unterricht. *Die Perspektive der Lehrkraft*
ISBN 978-3-8325-2006-9 43.00 EUR
- 87 Marita Schmidt: Kompetenzmodellierung und -diagnostik im Themengebiet Energie der Sekundarstufe I. *Entwicklung und Erprobung eines Testinventars*
ISBN 978-3-8325-2024-3 37.00 EUR
- 88 Gudrun Franke-Braun: Aufgaben mit gestuften Lernhilfen. *Ein Aufgabenformat zur Förderung der sachbezogenen Kommunikation und Lernleistung für den naturwissenschaftlichen Unterricht*
ISBN 978-3-8325-2026-7 38.00 EUR
- 89 Silke Klos: Kompetenzförderung im naturwissenschaftlichen Anfangsunterricht. *Der Einfluss eines integrierten Unterrichtskonzepts*
ISBN 978-3-8325-2133-2 37.00 EUR

- 90 Ulrike Elisabeth Burkard: Quantenphysik in der Schule. *Bestandsaufnahme, Perspektiven und Weiterentwicklungsmöglichkeiten durch die Implementation eines Medienservers*
ISBN 978-3-8325-2215-5 43.00 EUR
- 91 Ulrike Gromadecki: Argumente in physikalischen Kontexten. *Welche Geltungsgründe halten Physikanfänger für überzeugend?*
ISBN 978-3-8325-2250-6 41.50 EUR
- 92 Jürgen Bruns: Auf dem Weg zur Förderung naturwissenschaftsspezifischer Vorstellungen von zukünftigen Chemie-Lehrenden
ISBN 978-3-8325-2257-5 43.50 EUR
- 93 Cornelius Marsch: Räumliche Atomvorstellung. *Entwicklung und Erprobung eines Unterrichtskonzeptes mit Hilfe des Computers*
ISBN 978-3-8325-2293-3 82.50 EUR
- 94 Maja Brückmann: Sachstrukturen im Physikunterricht. *Ergebnisse einer Videostudie*
ISBN 978-3-8325-2272-8 39.50 EUR
- 95 Sabine Fechner: Effects of Context-oriented Learning on Student Interest and Achievement in Chemistry Education
ISBN 978-3-8325-2343-5 36.50 EUR
- 96 Clemens Nagel: eLearning im Physikalischen Anfängerpraktikum
ISBN 978-3-8325-2355-8 39.50 EUR
- 97 Josef Riese: Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften
ISBN 978-3-8325-2376-3 39.00 EUR
- 98 Sascha Bernholt: Kompetenzmodellierung in der Chemie. *Theoretische und empirische Reflexion am Beispiel des Modells hierarchischer Komplexität*
ISBN 978-3-8325-2447-0 40.00 EUR
- 99 Holger Christoph Stawitz: Auswirkung unterschiedlicher Aufgabenprofile auf die Schülerleistung. *Vergleich von Naturwissenschafts- und Problemlöseaufgaben der PISA 2003-Studie*
ISBN 978-3-8325-2451-7 37.50 EUR
- 100 Hans Ernst Fischer, Elke Sumfleth (Hrsg.): nwu-essen – 10 Jahre Essener Forschung zum naturwissenschaftlichen Unterricht
ISBN 978-3-8325-3331-1 40.00 EUR
- 101 Hendrik Härtig: Sachstrukturen von Physikschulbüchern als Grundlage zur Bestimmung der Inhaltsvalidität eines Tests
ISBN 978-3-8325-2512-5 34.00 EUR
- 102 Thomas Grüß-Niehaus: Zum Verständnis des Löslichkeitskonzeptes im Chemieunterricht. *Der Effekt von Methoden progressiver und kollaborativer Reflexion*
ISBN 978-3-8325-2537-8 40.50 EUR

- 103 Patrick Bronner: Quantenoptische Experimente als Grundlage eines Curriculums zur Quantenphysik des Photons
ISBN 978-3-8325-2540-8 36.00 EUR
- 104 Adrian Voßkühler: Blickbewegungsmessung an Versuchsaufbauten. *Studien zur Wahrnehmung, Verarbeitung und Usability von physikbezogenen Experimenten am Bildschirm und in der Realität*
ISBN 978-3-8325-2548-4 47.50 EUR
- 105 Verena Tobias: Newton'sche Mechanik im Anfangsunterricht. *Die Wirksamkeit einer Einführung über die zweidimensionale Dynamik auf das Lehren und Lernen*
ISBN 978-3-8325-2558-3 54.00 EUR
- 106 Christian Rogge: Entwicklung physikalischer Konzepte in aufgabenbasierten Lernumgebungen
ISBN 978-3-8325-2574-3 45.00 EUR
- 107 Mathias Ropohl: Modellierung von Schülerkompetenzen im Basiskonzept Chemische Reaktion. *Entwicklung und Analyse von Testaufgaben*
ISBN 978-3-8325-2609-2 36.50 EUR
- 108 Christoph Kulgemeyer: Physikalische Kommunikationskompetenz. *Modellierung und Diagnostik*
ISBN 978-3-8325-2674-0 44.50 EUR
- 109 Jennifer Olszewski: The Impact of Physics Teachers' Pedagogical Content Knowledge on Teacher Actions and Student Outcomes
ISBN 978-3-8325-2680-1 33.50 EUR
- 110 Annika Ohle: Primary School Teachers' Content Knowledge in Physics and its Impact on Teaching and Students' Achievement
ISBN 978-3-8325-2684-9 36.50 EUR
- 111 Susanne Mannel: Assessing scientific inquiry. *Development and evaluation of a test for the low-performing stage*
ISBN 978-3-8325-2761-7 40.00 EUR
- 112 Michael Plomer: Physik physiologisch passend praktiziert. *Eine Studie zur Lernwirksamkeit von traditionellen und adressatenspezifischen Physikpraktika für die Physiologie*
ISBN 978-3-8325-2804-1 34.50 EUR
- 113 Alexandra Schulz: Experimentierspezifische Qualitätsmerkmale im Chemieunterricht. *Eine Videostudie*
ISBN 978-3-8325-2817-1 40.00 EUR
- 114 Franz Boczianowski: Eine empirische Untersuchung zu Vektoren im Physikunterricht der Mittelstufe
ISBN 978-3-8325-2843-0 39.50 EUR
- 115 Maria Ploog: Internetbasiertes Lernen durch Textproduktion im Fach Physik
ISBN 978-3-8325-2853-9 39.50 EUR

- 116 Anja Dhein: Lernen in Explorier- und Experimentiersituationen. *Eine explorative Studie zu Bedeutungsentwicklungsprozessen bei Kindern im Alter zwischen 4 und 6 Jahren*
ISBN 978-3-8325-2859-1 45.50 EUR
- 117 Irene Neumann: Beyond Physics Content Knowledge. *Modeling Competence Regarding Nature of Scientific Inquiry and Nature of Scientific Knowledge*
ISBN 978-3-8325-2880-5 37.00 EUR
- 118 Markus Emden: Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. *Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I*
ISBN 978-3-8325-2867-6 38.00 EUR
- 119 Birgit Hofmann: Analyse von Blickbewegungen von Schülern beim Lesen von physikbezogenen Texten mit Bildern. *Eye Tracking als Methodenwerkzeug in der physikdidaktischen Forschung*
ISBN 978-3-8325-2925-3 59.00 EUR
- 120 Rebecca Knobloch: Analyse der fachinhaltlichen Qualität von Schüleräußerungen und deren Einfluss auf den Lernerfolg. *Eine Videostudie zu kooperativer Kleingruppenarbeit*
ISBN 978-3-8325-3006-8 36.50 EUR
- 121 Julia Hostenbach: Entwicklung und Prüfung eines Modells zur Beschreibung der Bewertungskompetenz im Chemieunterricht
ISBN 978-3-8325-3013-6 38.00 EUR
- 122 Anna Windt: Naturwissenschaftliches Experimentieren im Elementarbereich. *Evaluation verschiedener Lernsituationen*
ISBN 978-3-8325-3020-4 43.50 EUR
- 123 Eva Kölbach: Kontexteinflüsse beim Lernen mit Lösungsbeispielen
ISBN 978-3-8325-3025-9 38.50 EUR
- 124 Anna Lau: Passung und vertikale Vernetzung im Chemie- und Physikunterricht
ISBN 978-3-8325-3021-1 36.00 EUR
- 125 Jan Lamprecht: Ausbildungswege und Komponenten professioneller Handlungskompetenz. *Vergleich von Quereinsteigern mit Lehramtsabsolventen für Gymnasien im Fach Physik*
ISBN 978-3-8325-3035-8 38.50 EUR
- 126 Ulrike Böhm: Förderung von Verstehensprozessen unter Einsatz von Modellen
ISBN 978-3-8325-3042-6 41.00 EUR
- 127 Sabrina Dollny: Entwicklung und Evaluation eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Chemielehrkräften
ISBN 978-3-8325-3046-4 37.00 EUR
- 128 Monika Zimmermann: Naturwissenschaftliche Bildung im Kindergarten. *Eine integrative Längsschnittstudie zur Kompetenzentwicklung von Erzieherinnen*
ISBN 978-3-8325-3053-2 54.00 EUR

- 129 Ulf Saballus: Über das Schlussfolgern von Schülerinnen und Schülern zu öffentlichen Kontroversen mit naturwissenschaftlichem Hintergrund. *Eine Fallstudie*
ISBN 978-3-8325-3086-0 39.50 EUR
- 130 Olaf Krey: Zur Rolle der Mathematik in der Physik. *Wissenschaftstheoretische Aspekte und Vorstellungen Physiklernender*
ISBN 978-3-8325-3101-0 46.00 EUR
- 131 Angelika Wolf: Zusammenhänge zwischen der Eigenständigkeit im Physikunterricht, der Motivation, den Grundbedürfnissen und dem Lernerfolg von Schülern
ISBN 978-3-8325-3161-4 45.00 EUR
- 132 Johannes Börlin: Das Experiment als Lerngelegenheit. *Vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität*
ISBN 978-3-8325-3170-6 45.00 EUR
- 133 Olaf Uhden: Mathematisches Denken im Physikunterricht. *Theorieentwicklung und Problemanalyse*
ISBN 978-3-8325-3170-6 45.00 EUR
- 134 Christoph Gut: Modellierung und Messung experimenteller Kompetenz. *Analyse eines large-scale Experimentiertests*
ISBN 978-3-8325-3213-0 40.00 EUR
- 135 Antonio Rueda: Lernen mit ExploMultimedial in kolumbianischen Schulen. *Analyse von kurzzeitigen Lernprozessen und der Motivation beim länderübergreifenden Einsatz einer deutschen computergestützten multimedialen Lernumgebung für den naturwissenschaftlichen Unterricht*
ISBN 978-3-8325-3218-5 45.50 EUR
- 136 Krisztina Berger: Bilder, Animationen und Notizen. *Empirische Untersuchung zur Wirkung einfacher visueller Repräsentationen und Notizen auf den Wissenserwerb in der Optik*
ISBN 978-3-8325-3238-3 41.50 EUR
- 137 Antony Crossley: Untersuchung des Einflusses unterschiedlicher physikalischer Konzepte auf den Wissenserwerb in der Thermodynamik der Sekundarstufe I
ISBN 978-3-8325-3275-8 40.00 EUR
- 138 Tobias Viering: Entwicklung physikalischer Kompetenz in der Sekundarstufe I. *Validierung eines Kompetenzentwicklungsmodells für das Energiekonzept im Bereich Fachwissen*
ISBN 978-3-8325-3277-2 37.00 EUR
- 139 Nico Schreiber: Diagnostik experimenteller Kompetenz. *Validierung technologiegestützter Testverfahren im Rahmen eines Kompetenzstrukturmodells*
ISBN 978-3-8325-3284-0 39.00 EUR
- 140 Sarah Hundertmark: Einblicke in kollaborative Lernprozesse. *Eine Fallstudie zur reflektierenden Zusammenarbeit unterstützt durch die Methoden Concept Mapping und Lernbegleitbogen*
ISBN 978-3-8325-3251-2 43.00 EUR

- 141 Ronny Scherer: Analyse der Struktur, Messinvarianz und Ausprägung komplexer Problemlösekompetenz im Fach Chemie. *Eine Querschnittstudie in der Sekundarstufe I und am Übergang zur Sekundarstufe II*
ISBN 978-3-8325-3312-0 43.00 EUR
- 142 Patricia Heitmann: Bewertungskompetenz im Rahmen naturwissenschaftlicher Problemlöseprozesse. *Modellierung und Diagnose der Kompetenzen Bewertung und analytisches Problemlösen für das Fach Chemie*
ISBN 978-3-8325-3314-4 37.00 EUR
- 143 Jan Fleischhauer: Wissenschaftliches Argumentieren und Entwicklung von Konzepten beim Lernen von Physik
ISBN 978-3-8325-3325-0 35.00 EUR
- 144 Nermin Özcan: Zum Einfluss der Fachsprache auf die Leistung im Fach Chemie. *Eine Förderstudie zur Fachsprache im Chemieunterricht*
ISBN 978-3-8325-3328-1 36.50 EUR
- 145 Helena van Vorst: Kontextmerkmale und ihr Einfluss auf das Schülerinteresse im Fach Chemie
ISBN 978-3-8325-3321-2 38.50 EUR
- 146 Janine Cappell: Fachspezifische Diagnosekompetenz angehender Physiklehrkräfte in der ersten Ausbildungsphase
ISBN 978-3-8325-3356-4 38.50 EUR
- 147 Susanne Bley: Förderung von Transferprozessen im Chemieunterricht
ISBN 978-3-8325-3407-3 40.50 EUR
- 148 Cathrin Blaes: Die übungsgestützte Lehrerpräsentation im Chemieunterricht der Sekundarstufe I. *Evaluation der Effektivität*
ISBN 978-3-8325-3409-7 43.50 EUR
- 149 Julia Suckut: Die Wirksamkeit von piko-OWL als Lehrerfortbildung. Eine Evaluation zum Projekt *Physik im Kontext* in Fallstudien
ISBN 978-3-8325-3440-0 45.00 EUR
- 150 Alexandra Dorschu: Die Wirkung von Kontexten in Physikkompetenztestaufgaben
ISBN 978-3-8325-3446-2 37.00 EUR
- 151 Jochen Scheid: Multiple Repräsentationen, Verständnis physikalischer Experimente und kognitive Aktivierung: *Ein Beitrag zur Entwicklung der Aufgabenkultur*
ISBN 978-3-8325-3449-3 49.00 EUR
- 152 Tim Plasa: Die Wahrnehmung von Schülerlaboren und Schülerforschungszentren
ISBN 978-3-8325-3483-7 35.50 EUR
- 153 Felix Schoppmeier: Physikkompetenz in der gymnasialen Oberstufe. *Entwicklung und Validierung eines Kompetenzstrukturmodells für den Kompetenzbereich Umgang mit Fachwissen*
ISBN 978-3-8325-3502-5 36.00 EUR

- 154 Katharina Groß: Experimente alternativ dokumentieren. *Eine qualitative Studie zur Förderung der Diagnose- und Differenzierungskompetenz in der Chemielehrerbildung*
ISBN 978-3-8325-3508-7 43.50 EUR
- 155 Barbara Hank: Konzeptwandelprozesse im Anfangsunterricht Chemie. *Eine quasixperimentelle Längsschnittstudie*
ISBN 978-3-8325-3519-3 38.50 EUR
- 156 Katja Freyer: Zum Einfluss von Studieneingangsvoraussetzungen auf den Studienerfolg Erstsemesterstudierender im Fach Chemie
ISBN 978-3-8325-3544-5 38.00 EUR
- 157 Alexander Rachel: Auswirkungen instrukionaler Hilfen bei der Einführung des (Ferro-)Magnetismus. *Eine Vergleichsstudie in der Primar- und Sekundarstufe*
ISBN 978-3-8325-3548-3 43.50 EUR
- 158 Sebastian Ritter: Einfluss des Lerninhalts Nanogrößeneffekte auf Teilchen- und Teilchenmodellvorstellungen von Schülerinnen und Schülern
ISBN 978-3-8325-3558-2 36.00 EUR
- 159 Andrea Harbach: Problemorientierung und Vernetzung in kontextbasierten Lernaufgaben
ISBN 978-3-8325-3564-3 39.00 EUR
- 160 David Obst: Interaktive Tafeln im Physikunterricht. *Entwicklung und Evaluation einer Lehrerfortbildung*
ISBN 978-3-8325-3582-7 40.50 EUR
- 161 Sophie Kirschner: Modellierung und Analyse des Professionswissens von Physik Lehrkräften
ISBN 978-3-8325-3601-5 35.00 EUR
- 162 Katja Stief: Selbstregulationsprozesse und Hausaufgabenmotivation im Chemieunterricht
ISBN 978-3-8325-3631-2 34.00 EUR
- 163 Nicola Meschede: Professionelle Wahrnehmung der inhaltlichen Strukturierung im naturwissenschaftlichen Grundschulunterricht. *Theoretische Beschreibung und empirische Erfassung*
ISBN 978-3-8325-3668-8 37.00 EUR
- 164 Johannes Maximilian Barth: Experimentieren im Physikunterricht der gymnasialen Oberstufe. *Eine Rekonstruktion übergeordneter Einbettungsstrategien*
ISBN 978-3-8325-3681-7 39.00 EUR
- 165 Sandra Lein: Das Betriebspraktikum in der Lehrerbildung. *Eine Untersuchung zur Förderung der Wissenschafts- und Technikbildung im allgemeinbildenden Unterricht*
ISBN 978-3-8325-3698-5 40.00 EUR
- 166 Veranika Maiseyenka: Modellbasiertes Experimentieren im Unterricht. *Praxistauglichkeit und Lernwirkungen*
ISBN 978-3-8325-3708-1 38.00 EUR

- 167 Christoph Stolzenberger: Der Einfluss der didaktischen Lernumgebung auf das Erreichen geforderter Bildungsziele am Beispiel der W- und P-Seminare im Fach Physik
ISBN 978-3-8325-3708-1 38.00 EUR
- 168 Pia Altenburger: Mehrebenenregressionsanalysen zum Physiklernen im Sachunterricht der Primarstufe. *Ergebnisse einer Evaluationsstudie.*
ISBN 978-3-8325-3717-3 37.50 EUR
- 169 Nora Ferber: Entwicklung und Validierung eines Testinstruments zur Erfassung von Kompetenzentwicklung im Fach Chemie in der Sekundarstufe I
ISBN 978-3-8325-3727-2 39.50 EUR
- 170 Anita Stender: Unterrichtsplanung: Vom Wissen zum Handeln.
Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung
ISBN 978-3-8325-3750-0 41.50 EUR
- 171 Jenna Koenen: Entwicklung und Evaluation von experimentunterstützten Lösungsbeispielen zur Förderung naturwissenschaftlich-experimenteller Arbeitsweisen
ISBN 978-3-8325-3785-2 43.00 EUR
- 172 Teresa Henning: Empirische Untersuchung kontextorientierter Lernumgebungen in der Hochschuldidaktik. *Entwicklung und Evaluation kontextorientierter Aufgaben in der Studieneingangsphase für Fach- und Nebenfachstudierende der Physik*
ISBN 978-3-8325-3801-9 43.00 EUR
- 173 Alexander Pusch: Fachspezifische Instrumente zur Diagnose und individuellen Förderung von Lehramtsstudierenden der Physik
ISBN 978-3-8325-3829-3 38.00 EUR
- 174 Christoph Vogelsang: Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften. *Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz*
ISBN 978-3-8325-3846-0 50.50 EUR
- 175 Ingo Brebeck: Selbstreguliertes Lernen in der Studieneingangsphase im Fach Chemie
ISBN 978-3-8325-3859-0 37.00 EUR
- 176 Axel Eghtessad: Merkmale und Strukturen von Professionalisierungsprozessen in der ersten und zweiten Phase der Chemielehrerbildung. *Eine empirisch-qualitative Studie mit niedersächsischen Fachleiter_innen der Sekundarstufenlehrkräfte*
ISBN 978-3-8325-3861-3 45.00 EUR
- 177 Andreas Nehring: Wissenschaftliche Denk- und Arbeitsweisen im Fach Chemie. Eine kompetenzorientierte Modell- und Testentwicklung für den Bereich der Erkenntnisgewinnung
ISBN 978-3-8325-3872-9 39.50 EUR
- 178 Maike Schmidt: Professionswissen von Sachunterrichtslehrkräften. Zusammenhangsanalyse zur Wirkung von Ausbildungshintergrund und Unterrichtserfahrung auf das fachspezifische Professionswissen im Unterrichtsinhalt „Verbrennung“
ISBN 978-3-8325-3907-8 38.50 EUR

- 179 Jan Winkelmann: Auswirkungen auf den Fachwissenszuwachs und auf affektive Schülermerkmale durch Schüler- und Demonstrationsexperimente im Physikunterricht
ISBN 978-3-8325-3915-3 41.00 EUR
- 180 Iwen Kobow: Entwicklung und Validierung eines Testinstrumentes zur Erfassung der Kommunikationskompetenz im Fach Chemie
ISBN 978-3-8325-3927-6 34.50 EUR
- 181 Yvonne Gramzow: Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik. Modellierung und Testkonstruktion
ISBN 978-3-8325-3931-3 42.50 EUR
- 182 Evelin Schröter: Entwicklung der Kompetenzerwartung durch Lösen physikalischer Aufgaben einer multimedialen Lernumgebung
ISBN 978-3-8325-3975-7 54.50 EUR
- 183 Inga Kallweit: Effektivität des Einsatzes von Selbsteinschätzungsbögen im Chemieunterricht der Sekundarstufe I. *Individuelle Förderung durch selbstreguliertes Lernen*
ISBN 978-3-8325-3965-8 44.00 EUR
- 184 Andrea Schumacher: Paving the way towards authentic chemistry teaching. *A contribution to teachers' professional development*
ISBN 978-3-8325-3976-4 48.50 EUR
- 185 David Woitkowski: Fachliches Wissen Physik in der Hochschulausbildung. *Konzeptualisierung, Messung, Niveaubildung*
ISBN 978-3-8325-3988-7 53.00 EUR
- 186 Marianne Korner: Cross-Age Peer Tutoring in Physik. *Evaluation einer Unterrichtsmethode*
ISBN 978-3-8325-3979-5 38.50 EUR
- 187 Simone Nakoinz: Untersuchung zur Verknüpfung submikroskopischer und makroskopischer Konzepte im Fach Chemie
ISBN 978-3-8325-4057-9 38.50 EUR
- 188 Sandra Anus: Evaluation individueller Förderung im Chemieunterricht. *Adaptivität von Lerninhalten an das Vorwissen von Lernenden am Beispiel des Basiskonzeptes Chemische Reaktion*
ISBN 978-3-8325-4059-3 43.50 EUR
- 189 Thomas Roßbegalle: Fachdidaktische Entwicklungsforschung zum besseren Verständnis atmosphärischer Phänomene. *Treibhauseffekt, saurer Regen und stratosphärischer Ozonabbau als Kontexte zur Vermittlung von Basiskonzepten der Chemie*
ISBN 978-3-8325-4059-3 45.50 EUR
- 190 Kathrin Steckenmesser-Sander: Gemeinsamkeiten und Unterschiede physikbezogener Handlungs-, Denk- und Lernprozesse von Mädchen und Jungen
ISBN 978-3-8325-4066-1 38.50 EUR
- 191 Cornelia Geller: Lernprozessorientierte Sequenzierung des Physikunterrichts im Zusammenhang mit Fachwissenserwerb. *Eine Videostudie in Finnland, Deutschland und der Schweiz*
ISBN 978-3-8325-4082-1 35.50 EUR

- 192 Jan Hofmann: Untersuchung des Kompetenzaufbaus von Physiklehrkräften während einer Fortbildungsmaßnahme
ISBN 978-3-8325-4104-0 38.50 EUR
- 193 Andreas Dickhäuser: Chemiespezifischer Humor. *Theoriebildung, Materialentwicklung, Evaluation*
ISBN 978-3-8325-4108-8 37.00 EUR
- 194 Stefan Korte: Die Grenzen der Naturwissenschaft als Thema des Physikunterrichts
ISBN 978-3-8325-4112-5 57.50 EUR
- 195 Carolin Hülsmann: Kurswahlmotive im Fach Chemie. Eine Studie zum Wahlverhalten und Erfolg von Schülerinnen und Schülern in der gymnasialen Oberstufe
ISBN 978-3-8325-4144-6 49.00 EUR
- 196 Caroline Körbs: Mindeststandards im Fach Chemie am Ende der Pflichtschulzeit
ISBN 978-3-8325-4148-4 34.00 EUR
- 197 Andreas Vorholzer: Wie lassen sich Kompetenzen des experimentellen Denkens und Arbeitens fördern? *Eine empirische Untersuchung der Wirkung eines expliziten und eines impliziten Instruktionsansatzes*
ISBN 978-3-8325-4194-1 37.50 EUR
- 198 Anna Katharina Schmitt: Entwicklung und Evaluation einer Chemielehrerfortbildung zum Kompetenzbereich Erkenntnisgewinnung
ISBN 978-3-8325-4228-3 39.50 EUR
- 199 Christian Maurer: Strukturierung von Lehr-Lern-Sequenzen
ISBN 978-3-8325-4247-4 36.50 EUR
- 200 Helmut Fischler, Elke Sumfleth (Hrsg.): Professionelle Kompetenz von Lehrkräften der Chemie und Physik
ISBN 978-3-8325-4523-9 34.00 EUR
- 201 Simon Zander: Lehrerfortbildung zu Basismodellen und Zusammenhänge zum Fachwissen
ISBN 978-3-8325-4248-1 35.00 EUR
- 202 Kerstin Arndt: Experimentierkompetenz erfassen.
Analyse von Prozessen und Mustern am Beispiel von Lehramtsstudierenden der Chemie
ISBN 978-3-8325-4266-5 45.00 EUR
- 203 Christian Lang: Kompetenzorientierung im Rahmen experimentalchemischer Praktika
ISBN 978-3-8325-4268-9 42.50 EUR
- 204 Eva Cauet: Testen wir relevantes Wissen? *Zusammenhang zwischen dem Professionswissen von Physiklehrkräften und gutem und erfolgreichem Unterrichten*
ISBN 978-3-8325-4276-4 39.50 EUR
- 205 Patrick Löffler: Modellanwendung in Problemlöseaufgaben. *Wie wirkt Kontext?*
ISBN 978-3-8325-4303-7 35.00 EUR

- 206 Carina Gehlen: Kompetenzstruktur naturwissenschaftlicher Erkenntnisgewinnung im Fach Chemie
ISBN 978-3-8325-4318-1 43.00 EUR
- 207 Lars Oettinghaus: Lehrerüberzeugungen und physikbezogenes Professionswissen. *Vergleich von Absolventinnen und Absolventen verschiedener Ausbildungswege im Physikreferendariat*
ISBN 978-3-8325-4319-8 38.50 EUR
- 208 Jennifer Petersen: Zum Einfluss des Merkmals Humor auf die Gesundheitsförderung im Chemieunterricht der Sekundarstufe I. *Eine Interventionsstudie zum Thema Sonnenschutz*
ISBN 978-3-8325-4348-8 40.00 EUR
- 209 Philipp Straube: Modellierung und Erfassung von Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung bei (Lehramts-) Studierenden im Fach Physik
ISBN 978-3-8325-4351-8 35.50 EUR
- 210 Martin Dickmann: Messung von Experimentierfähigkeiten. *Validierungsstudien zur Qualität eines computerbasierten Testverfahrens*
ISBN 978-3-8325-4356-3 41.00 EUR
- 211 Markus Bohlmann: Science Education. Empirie, Kulturen und Mechanismen der Didaktik der Naturwissenschaften
ISBN 978-3-8325-4377-8 44.00 EUR
- 212 Martin Draude: Die Kompetenz von Physiklehrkräften, Schwierigkeiten von Schülerinnen und Schülern beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-4382-2 37.50 EUR
- 213 Henning Rode: Prototypen evidenzbasierten Physikunterrichts. *Zwei empirische Studien zum Einsatz von Feedback und Blackboxes in der Sekundarstufe*
ISBN 978-3-8325-4389-1 42.00 EUR
- 214 Jan-Henrik Kechel: Schülerschwierigkeiten beim eigenständigen Experimentieren. *Eine qualitative Studie am Beispiel einer Experimentieraufgabe zum Hooke'schen Gesetz*
ISBN 978-3-8325-4392-1 55.00 EUR
- 215 Katharina Fricke: Classroom Management and its Impact on Lesson Outcomes in Physics. *A multi-perspective comparison of teaching practices in primary and secondary schools*
ISBN 978-3-8325-4394-5 40.00 EUR
- 216 Hannes Sander: Orientierungen von Jugendlichen beim Urteilen und Entscheiden in Kontexten nachhaltiger Entwicklung. *Eine rekonstruktive Perspektive auf Bewertungskompetenz in der Didaktik der Naturwissenschaft*
ISBN 978-3-8325-4434-8 46.00 EUR
- 217 Inka Haak: Maßnahmen zur Unterstützung kognitiver und metakognitiver Prozesse in der Studieneingangsphase. *Eine Design-Based-Research-Studie zum universitären Lernzentrum Physiktreff*
ISBN 978-3-8325-4437-9 46.50 EUR

- 218 Martina Brandenburger: Was beeinflusst den Erfolg beim Problemlösen in der Physik?
Eine Untersuchung mit Studierenden
ISBN 978-3-8325-4409-6 42.50 EUR
- 219 Corinna Helms: Entwicklung und Evaluation eines Trainings zur Verbesserung der Erklärqualität von Schülerinnen und Schülern im Gruppenpuzzle
ISBN 978-3-8325-4454-6 42.50 EUR
- 220 Viktoria Rath: Diagnostische Kompetenz von angehenden Physiklehrkräften. *Modellierung, Testinstrumentenentwicklung und Erhebung der Performanz bei der Diagnose von Schülervorstellungen in der Mechanik*
ISBN 978-3-8325-4456-0 42.50 EUR
- 221 Janne Krüger: Schülerperspektiven auf die zeitliche Entwicklung der Naturwissenschaften
ISBN 978-3-8325-4457-7 45.50 EUR
- 222 Stefan Mutke: Das Professionswissen von Chemiereferendarinnen und -referendaren in Nordrhein-Westfalen. *Eine Längsschnittstudie*
ISBN 978-3-8325-4458-4 37.50 EUR
- 223 Sebastian Habig: Systematisch variierte Kontextaufgaben und ihr Einfluss auf kognitive und affektive Schülerfaktoren
ISBN 978-3-8325-4467-6 40.50 EUR
- 224 Sven Liepertz: Zusammenhang zwischen dem Professionswissen von Physiklehrkräften, dem sachstrukturellen Angebot des Unterrichts und der Schülerleistung
ISBN 978-3-8325-4480-5 34.00 EUR
- 225 Elina Platova: Optimierung eines Laborpraktikums durch kognitive Aktivierung
ISBN 978-3-8325-4481-2 39.00 EUR
- 226 Tim Reschke: Lesegeschichten im Chemieunterricht der Sekundarstufe I zur Unterstützung von situationalem Interesse und Lernerfolg
ISBN 978-3-8325-4487-4 41.00 EUR
- 227 Lena Mareike Walper: Entwicklung der physikbezogenen Interessen und selbstbezogenen Kognitionen von Schülerinnen und Schülern in der Übergangsphase von der Primar- in die Sekundarstufe. *Eine Längsschnittanalyse vom vierten bis zum siebten Schuljahr*
ISBN 978-3-8325-4495-9 43.00 EUR
- 228 Stefan Anthofer: Förderung des fachspezifischen Professionswissens von Chemielehramtsstudierenden
ISBN 978-3-8325-4498-0 39.50 EUR
- 229 Marcel Bullinger: Handlungsorientiertes Physiklernen mit instruierten Selbsterklärungen in der Primarstufe. *Eine experimentelle Laborstudie*
ISBN 978-3-8325-4504-8 44.00 EUR
- 230 Thomas Amenda: Bedeutung fachlicher Elementarisierungen für das Verständnis der Kinematik
ISBN 978-3-8325-4531-4 43.50 EUR

- 231 Sabrina Milke: Beeinflusst *Priming* das Physiklernen?
Eine empirische Studie zum Dritten Newtonschen Axiom
ISBN 978-3-8325-4549-4 42.00 EUR
- 232 Corinna Erfmann: Ein anschaulicher Weg zum Verständnis der elektromagnetischen Induktion. *Evaluation eines Unterrichtsvorschlags und Validierung eines Leistungsd Diagnoseinstruments*
ISBN 978-3-8325-4550-5 49.50 EUR
- 233 Hanne Rautenstrauch: Erhebung des (Fach-)Sprachstandes bei Lehramtsstudierenden im Kontext des Faches Chemie
ISBN 978-3-8325-4556-7 40.50 EUR
- 234 Tobias Klug: Wirkung kontextorientierter physikalischer Praktikumsversuche auf Lernprozesse von Studierenden der Medizin
ISBN 978-3-8325-4558-1 37.00 EUR
- 235 Mareike Bohrmann: Zur Förderung des Verständnisses der Variablenkontrolle im naturwissenschaftlichen Sachunterricht
ISBN 978-3-8325-4559-8 52.00 EUR
- 236 Anja Schödl: FALKO-Physik – Fachspezifische Lehrerkompetenzen im Fach Physik. *Entwicklung und Validierung eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Physiklehrkräften*
ISBN 978-3-8325-4553-6 40.50 EUR
- 237 Hilda Scheuermann: Entwicklung und Evaluation von Unterstützungsmaßnahmen zur Förderung der Variablenkontrollstrategie beim Planen von Experimenten
ISBN 978-3-8325-4568-0 39.00 EUR
- 238 Christian G. Strippel: Naturwissenschaftliche Erkenntnisgewinnung an chemischen Inhalten vermitteln. *Konzeption und empirische Untersuchung einer Ausstellung mit Experimentierstation*
ISBN 978-3-8325-4577-2 41.50 EUR
- 239 Sarah Rau: Durchführung von Sachunterricht im Vorbereitungsdienst. *Eine längsschnittliche, videobasierte Unterrichtsanalyse*
ISBN 978-3-8325-4579-6 46.00 EUR
- 240 Thomas Plotz: Lernprozesse zu nicht-sichtbarer Strahlung. *Empirische Untersuchungen in der Sekundarstufe 2*
ISBN 978-3-8325-4624-3 39.50 EUR
- 241 Wolfgang Aschauer: Elektrische und magnetische Felder. *Eine empirische Studie zu Lernprozessen in der Sekundarstufe II*
ISBN 978-3-8325-4625-0 50.00 EUR
- 242 Anna Donhauser: Didaktisch rekonstruierte Materialwissenschaft. *Aufbau und Konzeption eines Schülerlabors für den Exzellenzcluster Engineering of Advanced Materials*
ISBN 978-3-8325-4636-6 39.00 EUR

- 243 Katrin Schüller: Lernen mit Lösungsbeispielen im Chemieunterricht. *Einflüsse auf Lernerfolg, kognitive Belastung und Motivation*
ISBN 978-3-8325-4640-3 42.50 EUR
- 244 Timo Fleischer: Untersuchung der chemischen Fachsprache unter besonderer Berücksichtigung chemischer Repräsentationen
ISBN 978-3-8325-4642-7 46.50 EUR
- 245 Rosina Steininger: Concept Cartoons als Stimuli für Kleingruppendiskussionen im Chemieunterricht. *Beschreibung und Analyse einer komplexen Lerngelegenheit*
ISBN 978-3-8325-4647-2 39.00 EUR
- 246 Daniel Rehfeldt: Erfassung der Lehrqualität naturwissenschaftlicher Experimentalpraktika
ISBN 978-3-8325-4590-1 40.00 EUR
- 247 Sandra Puddu: Implementing Inquiry-based Learning in a Diverse Classroom: Investigating Strategies of Scaffolding and Students' Views of Scientific Inquiry
ISBN 978-3-8325-4591-8 35.50 EUR
- 248 Markus Bliersbach: Kreativität in der Chemie. *Erhebung und Förderung der Vorstellungen von Chemielehramtsstudierenden*
ISBN 978-3-8325-4593-2 44.00 EUR
- 249 Lennart Kimpel: Aufgaben in der Allgemeinen Chemie. *Zum Zusammenspiel von chemischem Verständnis und Rechenfähigkeit*
ISBN 978-3-8325-4618-2 36.00 EUR
- 250 Louise Bindel: Effects of integrated learning: explicating a mathematical concept in inquiry-based science camps
ISBN 978-3-8325-4655-7 37.50 EUR
- 251 Michael Wenzel: Computereinsatz in Schule und Schülerlabor. *Einstellung von Physiklehrkräften zu Neuen Medien*
ISBN 978-3-8325-4659-5 38.50 EUR
- 252 Laura Muth: Einfluss der Auswertephase von Experimenten im Physikunterricht. *Ergebnisse einer Interventionsstudie zum Zuwachs von Fachwissen und experimenteller Kompetenz von Schülerinnen und Schülern*
ISBN 978-3-8325-4675-5 36.50 EUR
- 253 Annika Fricke: Interaktive Skripte im Physikalischen Praktikum. *Entwicklung und Evaluation von Hypermedien für die Nebenfachausbildung*
ISBN 978-3-8325-4676-2 41.00 EUR
- 254 Julia Haase: Selbstbestimmtes Lernen im naturwissenschaftlichen Sachunterricht. *Eine empirische Interventionsstudie mit Fokus auf Feedback und Kompetenzerleben*
ISBN 978-3-8325-4685-4 38.50 EUR
- 255 Antje J. Heine: Was ist Theoretische Physik? *Eine wissenschaftstheoretische Betrachtung und Rekonstruktion von Vorstellungen von Studierenden und Dozenten über das Wesen der Theoretischen Physik*
ISBN 978-3-8325-4691-5 46.50 EUR

- 256 Claudia Meinhardt: Entwicklung und Validierung eines Testinstruments zu Selbstwirksamkeitserwartungen von (angehenden) Physiklehrkräften in physikdidaktischen Handlungsfeldern
ISBN 978-3-8325-4712-7 47.00 EUR
- 257 Ann-Kathrin Schlüter: Professionalisierung angehender Chemielehrkräfte für einen Gemeinsamen Unterricht
ISBN 978-3-8325-4713-4 53.50 EUR
- 258 Stefan Richtberg: Elektronenbahnen in Feldern. Konzeption und Evaluation einer webbasierten Lernumgebung
ISBN 978-3-8325-4723-3 49.00 EUR
- 259 Jan-Philipp Burde: Konzeption und Evaluation eines Unterrichtskonzepts zu einfachen Stromkreisen auf Basis des Elektronengasmodells
ISBN 978-3-8325-4726-4 57.50 EUR
- 260 Frank Finkenberg: Flipped Classroom im Physikunterricht
ISBN 978-3-8325-4737-4 42.50 EUR
- 261 Florian Treisch: Die Entwicklung der Professionellen Unterrichtswahrnehmung im Lehr-Lern-Labor Seminar
ISBN 978-3-8325-4741-4 41.50 EUR
- 262 Desiree Mayr: Strukturiertheit des experimentellen naturwissenschaftlichen Problemlöseprozesses
ISBN 978-3-8325-4757-8 37.00 EUR
- 263 Katrin Weber: Entwicklung und Validierung einer Learning Progression für das Konzept der chemischen Reaktion in der Sekundarstufe I
ISBN 978-3-8325-4762-2 48.50 EUR
- 264 Hauke Bartels: Entwicklung und Bewertung eines performanznahen Videovignetten-tests zur Messung der Erklärfähigkeit von Physiklehrkräften
ISBN 978-3-8325-4804-9 37.00 EUR
- 265 Karl Marniok: Zum Wesen von Theorien und Gesetzen in der Chemie. *Begriffsanalyse und Förderung der Vorstellungen von Lehramtsstudierenden*
ISBN 978-3-8325-4805-6 42.00 EUR
- 266 Marisa Holzapfel: Fachspezifischer Humor als Methode in der Gesundheitsbildung im Übergang von der Primarstufe zur Sekundarstufe I
ISBN 978-3-8325-4808-7 50.00 EUR
- 267 Anna Stolz: Die Auswirkungen von Experimentiersituationen mit unterschiedlichem Öffnungsgrad auf Leistung und Motivation der Schülerinnen und Schüler
ISBN 978-3-8325-4781-3 38.00 EUR
- 268 Nina Ulrich: Interaktive Lernaufgaben in dem digitalen Schulbuch eChemBook. *Einfluss des Interaktivitätsgrads der Lernaufgaben und des Vorwissens der Lernenden auf den Lernerfolg*
ISBN 978-3-8325-4814-8 43.50 EUR

- 269 Kim-Alessandro Weber: Quantenoptik in der Lehrerfortbildung. *Ein bedarfsgeprägtes Fortbildungskonzept zum Quantenobjekt Photon mit Realexperimenten*
ISBN 978-3-8325-4792-9 55.00 EUR
- 270 Nina Skorsetz: Empathisierer und Systematisierer im Vorschulalter. *Eine Fragebogen- und Videostudie zur Motivation, sich mit Naturphänomenen zu beschäftigen*
ISBN 978-3-8325-4825-4 43.50 EUR
- 271 Franziska Kehne: Analyse des Transfers von kontextualisiert erworbenem Wissen im Fach Chemie
ISBN 978-3-8325-4846-9 45.00 EUR
- 272 Markus Elsholz: Das akademische Selbstkonzept angehender Physiklehrkräfte als Teil ihrer professionellen Identität. *Dimensionalität und Veränderung während einer zentralen Praxisphase*
ISBN 978-3-8325-4857-5 37.50 EUR
- 273 Joachim Müller: Studienerfolg in der Physik. *Zusammenhang zwischen Modellierungskompetenz und Studienerfolg*
ISBN 978-3-8325-4859-9 35.00 EUR
- 274 Jennifer Dörschelln: Organische Leuchtdioden. *Implementation eines innovativen Themas in den Chemieunterricht*
ISBN 978-3-8325-4865-0 59.00 EUR
- 275 Stephanie Strelow: Beliefs von Studienanfängern des Kombi-Bachelors Physik über die Natur der Naturwissenschaften
ISBN 978-3-8325-4881-0 40.50 EUR
- 276 Dennis Jaeger: Kognitive Belastung und aufgabenspezifische sowie personenspezifische Einflussfaktoren beim Lösen von Physikaufgaben
ISBN 978-3-8325-4928-2 50.50 EUR
- 277 Vanessa Fischer: Der Einfluss von Interesse und Motivation auf die Messung von Fach- und Bewertungskompetenz im Fach Chemie
ISBN 978-3-8325-4933-6 39.00 EUR
- 278 René Dohrmann: Professionsbezogene Wirkungen einer Lehr-Lern-Labor-Veranstaltung. *Eine multimethodische Studie zu den professionsbezogenen Wirkungen einer Lehr-Lern-Labor-Blockveranstaltung auf Studierende der Bachelorstudiengänge Lehramt Physik und Grundschulpädagogik (Sachunterricht)*
ISBN 978-3-8325-4958-9 40.00 EUR
- 279 Meike Bergs: Can We Make Them Use These Strategies? *Fostering Inquiry-Based Science Learning Skills with Physical and Virtual Experimentation Environments*
ISBN 978-3-8325-4962-6 39.50 EUR
- 280 Marie-Therese Hauerstein: Untersuchung zur Effektivität von Strukturierung und Binnendifferenzierung im Chemieunterricht der Sekundarstufe I. *Evaluation der Strukturierungshilfe Lernleiter*
ISBN 978-3-8325-4982-4 42.50 EUR

- 281 Verena Zucker: Erkennen und Beschreiben von formativem Assessment im naturwissenschaftlichen Grundschulunterricht. *Entwicklung eines Instruments zur Erfassung von Teilfähigkeiten der professionellen Wahrnehmung von Lehramtsstudierenden*
ISBN 978-3-8325-4991-6 38.00 EUR
- 282 Victoria Telser: Erfassung und Förderung experimenteller Kompetenz von Lehrkräften im Fach Chemie
ISBN 978-3-8325-4996-1 50.50 EUR
- 283 Kristine Tschirschky: Entwicklung und Evaluation eines gedächtnisorientierten Aufgabendesigns für Physikaufgaben
ISBN 978-3-8325-5002-8 42.50 EUR
- 284 Thomas Elert: Course Success in the Undergraduate General Chemistry Lab
ISBN 978-3-8325-5004-2 41.50 EUR
- 285 Britta Kalthoff: Explizit oder implizit? *Untersuchung der Lernwirksamkeit verschiedener fachmethodischer Instruktionen im Hinblick auf fachmethodische und fachinhaltliche Fähigkeiten von Sachunterrichtsstudierenden*
ISBN 978-3-8325-5013-4 37.50 EUR
- 286 Thomas Dickmann: Visuelles Modellverständnis und Studienerfolg in der Chemie. *Zwei Seiten einer Medaille*
ISBN 978-3-8325-5016-5 44.00 EUR
- 287 Markus Sebastian Feser: Physiklehrkräfte korrigieren Schülertexte. *Eine Explorationsstudie zur fachlich-konzeptuellen und sprachlichen Leistungsfeststellung und -beurteilung im Physikunterricht*
ISBN 978-3-8325-5020-2 49.00 EUR
- 288 Matylda Dudzinska: Lernen mit Beispielaufgaben und Feedback im Physikunterricht der Sekundarstufe 1. *Energieerhaltung zur Lösung von Aufgaben nutzen*
ISBN 978-3-8325-5025-7 47.00 EUR
- 289 Ines Sonnenschein: Naturwissenschaftliche Denk- und Arbeitsprozesse Studierender im Labor
ISBN 978-3-8325-5033-2 52.00 EUR
- 290 Florian Simon: Der Einfluss von Betreuung und Betreuenden auf die Wirksamkeit von Schülerlaborbesuchen. *Eine Zusammenhangsanalyse von Betreuungsqualität, Betreuermerkmalen und Schülerlaborzielen sowie Replikationsstudie zur Wirksamkeit von Schülerlaborbesuchen*
ISBN 978-3-8325-5036-3 49.50 EUR
- 291 Marie-Annette Geyer: Physikalisch-mathematische Darstellungswechsel funktionaler Zusammenhänge. *Das Vorgehen von SchülerInnen der Sekundarstufe 1 und ihre Schwierigkeiten*
ISBN 978-3-8325-5047-9 46.50 EUR
- 292 Susanne Digel: Messung von Modellierungskompetenz in Physik. *Theoretische Herleitung und empirische Prüfung eines Kompetenzmodells physikspezifischer Modellierungskompetenz*
ISBN 978-3-8325-5055-4 41.00 EUR

- 293 Sönke Janssen: Angebots-Nutzungs-Prozesse eines Schülerlabors analysieren und gestalten. *Ein design-based research Projekt*
ISBN 978-3-8325-5065-3 57.50 EUR
- 294 Knut Wille: Der Productive Failure Ansatz als Beitrag zur Weiterentwicklung der Aufgabenkultur
ISBN 978-3-8325-5074-5 49.00 EUR
- 295 Lisanne Kraeva: Problemlösestrategien von Schülerinnen und Schülern diagnostizieren
ISBN 978-3-8325-5110-0 59.50 EUR
- 296 Jenny Lorentzen: Entwicklung und Evaluation eines Lernangebots im Lehramtsstudium Chemie zur Förderung von Vernetzungen innerhalb des fachbezogenen Professionswissens
ISBN 978-3-8325-5120-9 39.50 EUR
- 297 Micha Winkelmann: Lernprozesse in einem Schülerlabor unter Berücksichtigung individueller naturwissenschaftlicher Interessenstrukturen
ISBN 978-3-8325-5147-6 48.50 EUR
- 298 Carina Wöhlke: Entwicklung und Validierung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung angehender Physiklehrkräfte
ISBN 978-3-8325-5149-0 43.00 EUR
- 299 Thomas Schubatzky: Das Amalgam Anfangs-Elektrizitätslehreunterricht. *Eine multiperspektivische Betrachtung in Deutschland und Österreich*
ISBN 978-3-8325-5159-9 50.50 EUR
- 300 Amany Annaggar: A Design Framework for Video Game-Based Gamification Elements to Assess Problem-solving Competence in Chemistry Education
ISBN 978-3-8325-5150-6 52.00 EUR
- 301 Alexander Engl: CHEMIE PUR – Unterrichten in der Natur: *Entwicklung und Evaluation eines kontextorientierten Unterrichtskonzepts im Bereich Outdoor Education zur Veränderung der Einstellung zu „Chemie und Natur“*
ISBN 978-3-8325-5174-2 59.00 EUR
- 302 Christin Marie Sajons: Kognitive und motivationale Dynamik in Schülerlaboren. *Kontextualisierung, Problemorientierung und Autonomieunterstützung der didaktischen Struktur analysieren und weiterentwickeln*
ISBN 978-3-8325-5155-1 56.00 EUR
- 303 Philipp Bitzenbauer: Quantenoptik an Schulen. *Studie im Mixed-Methods Design zur Evaluation des Erlanger Unterrichtskonzepts zur Quantenoptik*
ISBN 978-3-8325-5123-0 59.00 EUR
- 304 Malte S. Ubben: Typisierung des Verständnisses mentaler Modelle mittels empirischer Datenerhebung am Beispiel der Quantenphysik
ISBN 978-3-8325-5181-0 43.50 EUR
- 305 Wiebke Kuske-Janßen: Sprachlicher Umgang mit Formeln von LehrerInnen im Physikunterricht am Beispiel des elektrischen Widerstandes in Klassenstufe 8
ISBN 978-3-8325-5183-4 47.50 EUR

- 306 Kai Bliesmer: Physik der Küste für außerschulische Lernorte. *Eine Didaktische Rekonstruktion*
ISBN 978-3-8325-5190-2 58.00 EUR
- 307 Nikola Schild: Eignung von domänenspezifischen Studieneingangsvariablen als Prädiktoren für Studienerfolg im Fach und Lehramt Physik
ISBN 978-3-8325-5226-8 42.00 EUR
- 308 Daniel Averbeck: Zum Studienerfolg in der Studieneingangsphase des Chemiestudiums. *Der Einfluss kognitiver und affektiv-motivationaler Variablen*
ISBN 978-3-8325-5227-5 51.00 EUR
- 309 Martina Strübe: Modelle und Experimente im Chemieunterricht. *Eine Videostudie zum fachspezifischen Lehrerwissen und -handeln*
ISBN 978-3-8325-5245-9 45.50 EUR
- 310 Wolfgang Becker: Auswirkungen unterschiedlicher experimenteller Repräsentationen auf den Kenntnisstand bei Grundschulkindern
ISBN 978-3-8325-5255-8 50.00 EUR
- 311 Marvin Rost: 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
- 312 Christina Kobl: Förderung und Erfassung der Reflexionskompetenz im Fach Chemie
ISBN 978-3-8325-5259-6 41.00 EUR

Alle erschienenen Bücher können unter der angegebenen ISBN direkt online (<http://www.logos-verlag.de>) oder per Fax (030 - 42 85 10 92) beim Logos Verlag Berlin bestellt werden.

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth

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. Hans Niedderer
Inst. f. Didaktik der Naturwissenschaften,
Abt. Physikdidaktik,
FB Physik/ Elektrotechnik,
Universität Bremen,
Postfach 33 04 40, 28334 Bremen
Tel. 0421-218 4695 (Sekretariat),
e-mail: niedderer@physik.uni-bremen.de

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

Prof. Dr. Elke Sumfleth
Didaktik der Chemie,
Fakultät für Chemie,
Universität Duisburg-Essen,
Schützenbahn 70, 45127 Essen
Tel. 0201-183 3757/3761,
e-mail: elke.sumfleth@uni-due.de

Reflexionskompetenz kann als wichtige Voraussetzung für die Professionalisierung von Lehrkräften erachtet werden. Weder sind die Förderung noch der Einfluss von Reflexionskompetenz auf Unterrichtsplanungen im Fach Chemie ausreichend erforscht. Um diesem Desiderat zu begegnen, wurde ein universitäres Seminar konzipiert und evaluiert.

Wesentliche Ziele waren die Untersuchung möglicher Fördermaßnahmen zur Ausbildung der Reflexionskompetenz sowie die Klärung des Zusammenhangs zur Entwicklung des Fachwissens, des fachdidaktischen Wissens und der Selbstwirksamkeitserwartung von Chemielehramtsstudierenden.

Im Rahmen eines Pre-Post-Follow Up Kontrollgruppendesigns wurden Unterrichtskonzepte entwickelt und diese an drei Schülertagen erprobt. Im Anschluss wurden jeweils die Videos der Stunden schriftlich reflektiert. Die zwei Interventionsgruppen (mit/ohne ergänzendem externen Feedback) konnten ihre überarbeiteten Konzepte wiederholt erproben, während die Kontrollgruppe lediglich hospitierete. Die Entwicklung der Reflexionskompetenz wurde mithilfe eines Wissenstests und eines Kodiermanuals anhand der schriftlichen Unterrichtsreflexionen erfasst. Während die Reflexionskompetenz in beiden Interventionsgruppen signifikant gesteigert werden konnte, wurden Zusammenhänge mit dem fachdidaktischen Wissen und dem Lernerfolg der Schülerinnen und Schüler nur teilweise gefunden. Allgemein scheint die Fördermaßnahme Feedback einen eher geringen Einfluss auf die Ausbildung der Reflexionskompetenz im untersuchten Themenfeld zu haben.

Logos Verlag Berlin

ISBN 978-3-8325-5259-6