pedocs
Open Access Erziehungswissenschaften

DIPF

Albrecht, Luc Broder

Kritisches Denken in der Chemielehrer*innenbildung. Eine konzeptionell-empirische Studie zur Integration und Operationalisierung

Berlin : Logos Verlag Berlin 2025, 455 S. - (Studien zum Physik- und Chemielernen; 391)

Quellenangabe/ Reference:
Albrecht, Luc Broder: Kritisches Denken in der Chemielehrer*innenbildung. Eine konzeptionell-empirische Studie zur Integration und Operationalisierung. Berlin : Logos Verlag Berlin 2025, 455 S. - (Studien zum Physik- und Chemielernen; 391) - URN: urn:nbn:de:0111-pedocs-349968 - DOI: 10.25656/01:34996; https://doi.org/10.30819/5970; 10.30819/5970

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-349968
https://doi.org/10.25656/01:34996

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 und M. Ropohl [Hrsg.]

391

Luc Broder Albrecht

Kritisches Denken in der Chemielehrer*innenbildung

Eine konzeptionell-empirische Studie zur Integration und Operationalisierung

λογος

Studien zum Physik- und Chemielernen

Band 391

Luc Broder Albrecht

Kritisches Denken in der Chemielehrer*innenbildung

Eine konzeptionell-empirische Studie zur Integration und Operationalisierung

Logos Verlag Berlin

Studien zum Physik- und Chemielernen
Martin Hopf und Mathias Ropohl [Hrsg.]

Bibliografische Information der Deutschen Nationalbibliothek

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

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

Logos Verlag Berlin GmbH 2025

ISBN 978-3-8325-5970-0
ISSN 1614-8967
DOI 10.30819/5970

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

Kritisches Denken in der Chemielehrer*innenbildung

Eine konzeptionell-empirische Studie
zur Integration und Operationalisierung

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Luc Broder Albrecht

aus Köln

Köln 2025

Die vorliegende Dissertation wurde von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität zu Köln angenommen.

Berichterstatter/in:	Prof. Dr. Christiane S. Reiners
	Prof. Dr. Katharina Groß
Tag der mündlichen Prüfung:	02.05.2025

Inhaltsverzeichnis

Abbildungsverzeichnis VII

Tabellenverzeichnis XI

Abkürzungsverzeichnis XXI

Kurzzusammenfassung XXIII

Abstract XXV

1 Einleitung 1

2 Theoretische Grundlegungen 5

2.1 Die Überführung von Wissen in Handeln 5

2.1.1 Ziele naturwissenschaftlicher Grundbildung 5

2.1.2 Die Bedeutung von Nature of Science 12

2.1.3 Die Herausforderung durch die Medien 17

2.1.4 Pseudowissenschaften in der Gesellschaft 19

2.1.5 Das Abgrenzungsproblem 22

2.1.6 Gesellschaftlich relevante Probleme im Fokus 27

2.2 Kritisches Denken 31

2.2.1 Naturwissenschaftsdidaktisches Verständnis und Relevanz im naturwissenschaftlichen Unterricht 33

2.2.2 Die historische Entwicklung des Kritikbegriffs 40

2.2.3 Philosophische Perspektive 41

2.2.4 Psychologische Perspektive 45

2.2.5 Bildungswissenschaftliche Perspektive 47

2.3 Die Bedeutung der affektiven Ebene für das menschliche Denken 50

2.3.1 Erste Ansätze zum Verständnis menschlichen Handelns 51

2.3.2 Die Dual-Process Theory 51

2.3.3 Intuition und kognitive Verzerrungen 54

2.3.4 Genauere Betrachtung einzelner Einflussfaktoren 55

2.3.5 Theory of Planned Behaviour 57

2.3.6 Der Einfluss der affektiven Ebene auf die Anwendung von Wissen 59

3 Konzeptioneller Teil - Synthese 61

3.1 Ableitung der Forschungsfragen 61

3.2 Die Bedeutung der affektiven Ebene in den Naturwissenschaften 62

3.3 Die Definition Kritischen Denkens 67

3.4 Abgrenzung zu anderen Konzepten 74

3.5 Kritisches Denken im naturwissenschaftlichen Unterricht 76

4 Empirischer Teil – Methodik 87

4.1 Konzeption des empirischen Forschungsprojekts 87

4.2 Methoden der Datenerhebung 90

4.2.1 Datenschutz 93

4.2.2 Stichprobenbeschreibung der Erhebungen 95

4.2.3 Fragebogenerhebungen 100

4.2.4 Interviewerhebungen 106

4.2.5 Vermittlungseinheit 113

4.2.6 Verwendete Kontexte 119

4.2.6.1 Initialkontext: Handykauf 120
4.2.6.2 Interviewkontext A: Solarenergie vs. Windkraft 122
4.2.6.3 Interviewkontext B: E-Mobilität vs. Brennstoffzellenantrieb 126
4.2.6.4 Interviewkontext C: Plastik 130
4.2.6.5 Interviewkontext D: Fracking 132
4.2.6.6 Interviewkontext E: Stickoxidgrenzwerte 133
4.2.6.7 Zusatzkontext 1: Nanotechnologie 135
4.2.6.8 Zusatzkontext 2: Chemie in der Landwirtschaft 136
4.2.6.9 Zusatzkontext 3: Atomenergie 138

4.3 Methoden der Datenauswertung 139

4.3.1 Qualitative Auswertungsmethoden 139

4.3.2 Quantitative Auswertungsmethoden 151

4.3.2.1 Maße der zentralen Tendenz 151
4.3.2.2 Streumaße 152
4.3.2.3 Vergleichende Testverfahren 153

5 Beschreibung der Studien 163

5.1 Studie S1 164

5.1.1 S1.F1 - Fragengenese 1 165

5.1.1.1 Auswertung ... 166
5.1.1.2 Ergebnisse ... 169
5.1.2 S1.F2 - Fragengenese 2 ... 173
5.1.2.1 Auswertung ... 174
5.1.2.2 Ergebnisse ... 175
5.1.3 S1.F3 – Kriterien, Verständnis und CRT ... 178
5.1.3.1 Auswertung ... 180
5.1.3.2 Ergebnisse ... 185
5.1.4 Vermittlungseinheit ... 191
5.1.5 Abschluss der ersten Studie ... 194
5.2 STUDIE S2 ... 195
5.2.1 Interviewerhebung S2.I1 ... 196
5.2.1.1 Auswertung ... 198
5.2.1.2 Ergebnisse ... 203
5.2.2 Fragebogenerhebung S2.F1 ... 209
5.2.2.1 Auswertung ... 210
5.2.2.2 Ergebnisse ... 210
5.2.3 Abschluss der Studie S2 ... 212
5.3 STUDIE S3 ... 215
5.3.1 Interviewerhebung S3.I1&2 ... 217
5.3.1.1 Auswertung ... 218
5.3.1.2 Ergebnisse ... 229
5.3.2 Fragebogenerhebung S3.F1&2 ... 245
5.3.2.1 Auswertung ... 250
5.3.2.2 Ergebnisse ... 254
5.3.3 Interkoderreliabilität ... 272
5.3.4 Rückmeldungen der Teilnehmenden zur Vermittlungseinheit ... 274
5.3.5 Abschluss der Studie S3 ... 279
5.4 STUDIE S4 ... 281
5.4.1 Auswertung ... 283
5.4.2 Ergebnisse ... 284
5.4.3 Interkoderreliabilität ... 290
5.4.4 Abschluss der Studie S4 ... 291
5.5 STUDIE S5 ... 292
5.5.1 Interviewerhebung S5.I1&2 ... 293
5.5.2 Fragebogenerhebung ... 308

5.5.3 Interkoderreliabilität 329
5.5.4 Rückmeldungen der Teilnehmenden zur Vermittlungseinheit 331
5.5.5 Abschluss der Studie S5 334
5.6 Studienübergreifende Gruppenvergleiche 336
5.7 Zusammenführung der Ergebnisse 343
5.8 Abschließende Diskussion der durchgeführten Studien 351
6 Fazit und Ausblick 353
6.1 Beantwortung der konzeptionellen Forschungsfrage 353
6.2 Beantwortung der empirischen Forschungsfrage 357
6.3 Chemiedidaktische Implikationen und Ausblick 361
6.4 Ausblick auf zukünftige Forschungen 364
7 Zusammenfassung 367
Literaturverzeichnis 373
Anhang 405
Danksagung 453
Teilpublikationen 455

Abbildungsverzeichnis

Abbildung 1: The nine major sub-elements or key NOS aspects (McComas, 2020, S. 40) 14

Abbildung 2: Nature of Science-in-Society (NOSIS) (Höttecke & Allchin, 2020, S. 5) 16

Abbildung 3: Wege der Informationsverbreitung (Höttecke & Allchin, 2020, S. 7) 18

Abbildung 4: Kontinuum der Fachbereiche, welche das Konstrukt Kritischen Denkens berühren (eigene Darstellung) 33

Abbildung 5: Komponenten selbstregulierenden Lernens nach Schraw et al. (2006, S. 113) 35

Abbildung 6: Rahmen zur Förderung des Kritischen Denkens von Lernenden im naturwissenschaftlichen Unterricht (Vieira et al., 2011, S. 50) 36

Abbildung 7: CT-NOS Framework (Yacoubian, 2015, S. 253) 37

Abbildung 8: Synergiemodell des kritischen Denkens nach Rafolt et al. (2019, S. 66) 39

Abbildung 9: Integrative framework for critical thinking (Dwyer et al., 2014, S. 49) 46

Abbildung 10: Theorie des geplanten Verhaltens (Ajzen, 1991, S. 182) 58

Abbildung 11: Dual-Process Model of Critical Thinking (DP-CT) (eigenes Modell) 70

Abbildung 12: Kritisches Denken als Fundament und Vermittlerinstanz zwischen kognitiver und affektiver Dimension (eigene Darstellung) 77

Abbildung 13: Triadisches Modell Kritischen Denkens in der Naturwissenschaftsdidaktik (eigenes Modell) 81

Abbildung 14: Ebenenmodell des Kritischen Denkens in der Naturwissenschaftsdidaktik (eigene Darstellung) 83

Abbildung 15: Paralleles Mixed-Methods-Design (QUAL + QUAN) (eigene Darstellung basierend auf Hagenauer & Gläser-Zikuda, 2022, S. 256; in Anlehnung an Creswell, 2015, S. 56; Teddlie & Tashakkori, 2009, S. 152) 91

Abbildung 16: Überblick über den Aufbau des Forschungsprojekts. Entwicklung eines Instruments zur Darstellung der Dimensionen des Kritischen Denkens (eigene Darstellung in Anlehnung an Schumacher 2015, S. 53) 92

Abbildung 17: Ablauf der durchgeführten Kategorienbildung in Anlehnung an Kuckartz (2016) (eigene Darstellung) 144

Abbildung 18: Übersicht über den Aufbau und die Erhebungsparameter der Studie S1 164

Abbildung 19: Übersicht über den Aufbau und die Erhebungsparameter der Studie S2 195

Abbildung 20: Übersicht über den Aufbau und die Erhebungsparameter der Studie S3 216

Abbildung 21: Triadisches Modell Kritischen Denkens in der Naturwissenschaftsdidaktik, näher erläutert in Kapitel 3.5 218

Abbildung 22: S3.I1&2 - Gesamtübersicht über die prozentualen Anteile an den einzelnen Teilbereichen (% max. Kat. bezeichnet den prozentualen Anteil an den maximal möglichen Kategorien) 242

Abbildung 23: S3.I1&2 - Exemplarische Auswahl individueller Vergleichsdiagramme (% max. Kat. bezeichnet den prozentualen Anteil an der maximalen Anzahl möglicher Kategorien) 243

Abbildung 24: S3.I1&2 + FU - Gesamtübersicht und Vergleichsdiagramme der Teilnehmenden: Prozentuale Anteile an den einzelnen Teilbereichen (% max. Kat. bezeichnet den prozentualen Anteil an der maximalen Anzahl möglicher Kategorien) 245

Abbildung 25: Übersicht über den Aufbau und die Erhebungsparameter der Studie S4 281

Abbildung 26: Übersicht über den Aufbau und die Erhebungsparameter der Studie S5 292

Abbildung 27: S5.I1&2 - Gesamtübersicht über die prozentualen Anteile an den einzelnen Teilbereichen (% max. Kat. bezeichnet den prozentualen Anteil an der maximalen Anzahl möglicher Kategorien) 304

Abbildung 28: S5.I1&2 - Exemplarische Auswahl individueller Vergleichsdiagramme (% max. Kat. bezeichnet den prozentualen Anteil an den maximal möglichen Kategorien) 305

Abbildung 29: S5.I1&2 + Follow-Up - Gesamtübersicht über die prozentualen Anteile an den einzelnen Teilbereichen (% max. Kat. bezeichnet den prozentualen Anteil an der maximalen Anzahl möglicher Kategorien) 306

Abbildung 30: S5.I1&2 + FU - S5.I1&2 - Exemplarische Auswahl individueller Vergleichsdiagramme (% max. Kat. bezeichnet den prozentualen Anteil an der maximalen Anzahl möglicher Kategorien) 307

Abbildung 31: Übertragbarkeit der Ergebnisse auf das Triadische Modell Kritischen Denkens 350

Abbildung 32: Dual-Process Model of Critical Thinking (DP-CT) (eigenes Modell) 354

Abbildung 33: Triadisches Modell Kritischen Denkens in der Naturwissenschaftsdidaktik (eigenes Modell) 356

Abbildung 34: Netzdiagramm der Ergebnisse der Studie S5.I1&2 + Follow-Up 359

Tabellenverzeichnis

Tabelle 1: Übersicht über die Kern-Elemente von NOS. Gegenüberstellung von Lederman, McComas & Olson und Osborne et al. (Albrecht, 2021; vgl. Neumann & Kremer, 2013) 13

Tabelle 2: Ergebnisse der Delphi-Studie von Facione (1990). Kognitive Fähigkeiten und affektive Dispositionen von Kritischem Denken (eigene Übersetzung und Darstellung) 49

Tabelle 3: General Features of the Two Systems (Haidt, 2008, S. 1029) 52

Tabelle 4: Übersicht über die einzelnen Studien des Forschungsprojekts und die einzelnen Stichproben 93

Tabelle 5: Tabellarische Übersicht über die einzelnen Stichproben der Forschungsprojektes und ihrer Charakteristika 99

Tabelle 6: Einschätzung der Varianz von Rating-Skalen (Iacobucci & Churchill, 2010) 102

Tabelle 7: Items der SRIS nach Silvia (2021), basierend auf dem SRIS von Grant et al. (2002), die im späteren Verlauf rekodierten inversen Items sind durch (*) gekennzeichnet 102

Tabelle 8: Fragen des CRT nach Frederick (2005) 104

Tabelle 9: Vergleichswerte für die Auswertung des CRT (Frederick, 2005, S. 29) 105

Tabelle 10: Schematische Darstellung des Interviewleitfadens von Studie S2 110

Tabelle 11: Schematische Darstellung des Interviewleitfadens ab Studie S3 112

Tabelle 12: Übersicht über die Seminarsitzungen im SoSe23, Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP) 115

Tabelle 13: Auswählbare Kontexte für die Systemisch-reflexive Stellungnahme .. 116

Tabelle 14: Übersicht über die Seminarinhalte in Studie S5 118

Tabelle 15: Kritischen Werte für den Wilcoxon-Test, aus Bortz und Schuster, 2010, zitiert nach Clauss und Ebner, 1971, ebenfalls zu finden in Rasch et al., 2014a, S. 164 155

Tabelle 16: Kritischen Werte bei zweiseitigem Test bei $\alpha = 0{,}05$, aus Bortz & Schuster, 2010, S. 599; ebenfalls zu finden in Rasch et al., 2014a, S. 162 159

Tabelle 17: Chi-Quadrat-Verteilung für den Kurskal-Wallis-Test, aus Rasch et al., 2014a, S. 165; zitiert nach Hays & Winkler, 1970, S. 604–605 161

Tabelle 18: S1.F1 - Fragestellungen zur Fragengenese 166

Tabelle 19: S1.F1 - Übergeordnete Kategorien der Fragengenese 1 167

Tabelle 20: S1.F1 - Kategorien des *kriterialen Ansatzes* (KA) 168

Tabelle 21: S1.F1 - Kategorien des *reflexiven Ansatzes* (RA) 168

Tabelle 22: S1.F1 - Kategorien des *subjektiven Ansatzes* (SA) 169

Tabelle 23: S1.F1 - Ergebnisse der Fragebogenerhebung, Kategorisierung der Fragen, *N* = 15 170

Tabelle 24: S1.F1 - Ergebnisse der Fragebogenerhebung, Priorisierung der Fragen, *N* = 15 172

Tabelle 25: S1.F1 - Fragestellungen zur Fragengenese 174

Tabelle 26: S1.F2 - Kategorien des *kriterialen Ansatzes* (KA) 174

Tabelle 27: S1.F2 - Kategorien des *reflexiven Ansatzes* (RA) 175

Tabelle 28: S1.F2 - Ergebnisse der Fragebogenerhebung, Kategorisierung der Fragen, *N* = 12 176

Tabelle 29: S1.F2 - Ergebnisse der Fragebogenerhebung, Priorisierung der Fragen, *N* = 12 177

Tabelle 30: S1.F3 - Fragestellungen in der Fragebogenerhebung 178

Tabelle 31: S1.F3 - Untersuchungsfragen zu den Fragestellungen 180

Tabelle 32: S1F3 - Kategorien der Kriterien bei der Bewertung naturwissenschaftlicher Informationen in Medien (KB) 181

Tabelle 33: S1.F3 - Kategorien und Subkategorien von *Strategien bei der Bewertung naturwissenschaftlicher Informationen in Medien* (SB) 182

Tabelle 34: S1.F3 - Kategorien und Subkategorien des *Verständnisses naturwissenschaftlicher Arbeits- und Denkweisen* (AD) 183

Tabelle 35: S1.F3 - Kategorien und Subkategorien des *Verständnisses Kritischen Denkens* (KD) 184

Tabelle 36: S1.F3 - Ergebnisse der Fragebogenerhebung, *Kriterien bei der Bewertung naturwissenschaftlicher Informationen in Medien* (KB), *N* = 15 186

Tabelle 37: S1.F3 - Ergebnisse der Fragebogenerhebung, *Strategien bei der Bewertung naturwissenschaftlicher Informationen in Medien* (SB), *N* = 15 187

Tabelle 38: S1.F3 - Ergebnisse der Fragebogenerhebung, *Verständnis naturwissenschaftlicher Arbeits- und Denkweisen* (AD), *N* = 15 188

Tabelle 39: S1.F3 - Ergebnisse der Fragebogenerhebung, *Verständnis Kritisches Denken*, *N* = 15 190

Tabelle 40: S1.F3 - Ergebnisse der Fragebogenerhebung, CRT, *N* = 15 191

Tabelle 41: S1 - Übersicht über die Seminarsitzungen im SoSe22, Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP) 193

Tabelle 42: S2.I1 - Kategorien und Subkategorien der *Selbstreflexion bei der Bewertung* gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (SR) 199

Tabelle 43: S2.I1 - Kategorien des *Verständnisses von kritisch sein* (VK) 201

Tabelle 44: S2.I1 - Kategorien des *Verständnisses von Selbstreflexion* (VS) 202

Tabelle 45: S2.I1 - Kategorien zur *Einschätzung des Fachwissens* (FW) 202

Tabelle 46: S2.I1 - Kategorien zur *Einschätzung der Bewertungsfähigkeit* (BF) 203

Tabelle 47: S2.I1 - Ergebnisse der Interviewerhebung, Selbstreflexion bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (SR), *N* = 10 204

Tabelle 48: S2.I1 - Ergebnisse der Interviewerhebung, *Verständnis der Begrifflichkeit kritisch sein* (VK), *N* = 10 206

Tabelle 49: S2.I1 - Ergebnisse der Interviewerhebung, *Verständnis des Begriffs Selbstreflexion* (VS), *N* = 10 207

Tabelle 50: S2.I1 - Ergebnisse der Interviewerhebung, *Selbsteinschätzung des Fachwissens* (FW), *N* = 10 208

Tabelle 51: S2.I1 - Ergebnisse der Interviewerhebung, *Selbsteinschätzung der Bewertungsfähigkeit* (BF), *N* = 10 208

Tabelle 52: S2.F1 - Ergebnisse der SRIS nach Silvia (2021), *N* = 10 211

Tabelle 53: S2.F1 - Ergebnisse der SRIS nach Silvia (2021), Mittelwerte pro Proband*in, *N* = 10 211

Tabelle 54: S2.F1 - Ergebnisse des CRT nach Frederick (2005), *N* = 10 212

Tabelle 55: S3.I1&2 - Konzeptionelle Teilbereiche und zugehörige Untersuchungsfragen 219

Tabelle 56: S3.I1&2 - Kategorien der *Aspekte naturwissenschaftlicher Bewertung* (NW-A) 220

Tabelle 57: S3.I1&2 - Kategorien der *fachwissenschaftlichen Bewertung* (NW-FB) 221

Tabelle 58: S3.I1&2 - Kategorien der *Selbsteinschätzung des Fachwissens* (NW-SF) ... 222

Tabelle 59: S3.I1&2 - Kategorien der *Aspekte der Selbstreflexion* (SR-A) 223

Tabelle 60: S3.I1&2 - Kategorien des *Verständnisses der Selbstreflexion* (SR-V) 224

Tabelle 61: S3.I1&2 - Kategorien der *Aspekte des epistemologischen Verständnisses* (EP-A) 226

Tabelle 62: S3.I1&2 - Kategorien des *Verständnisses von Kritischem Denken* (KD-V) 227

Tabelle 63: S3.I1&2 - Kategorien der *Fähigkeit zur Bewertung* (FB) 228

Tabelle 64: S3.I1&2 - Ergebnisse der Interviewerhebung, *naturwissenschaftliche Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (NW), *N* = 10 230

Tabelle 65: S3.I1&2 - Ergebnisse der Interviewerhebung, *selbstreflexive Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (NW), *N* = 10 231

Tabelle 66: S3.I1&2 - Ergebnisse der Interviewerhebung, *epistemologische Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (NW), *N* = 10 233

Tabelle 67: S3.I1&2 - Ergebnisse der Interviewerhebung, *Verständnis Selbstreflexion* (SR-V), *N* = 10 235

Tabelle 68: S3.I1&2 - Ergebnisse der Interviewerhebung, *Verständnis Kritisches Denken* (KD-V), *N* = 10 237

Tabelle 69: S3.I1&2 - Ergebnisse der Interviewerhebung, *Selbsteinschätzung Fachwissen (NW-SF) und Fachwissenschaftliche Bewertung (NW-FB*), *N* = 10 238

Tabelle 70: S3.I1&2 - Ergebnisse der Interviewerhebung, *Fähigkeit zur Bewertung (FB)*, *N* = 10 239

Tabelle 71: S3.I1&2 - Ergebnisse der Interviewerhebung, Übersicht über die Einzelzählungen der verschiedenen Teilbereiche, *N* = 10 240

Tabelle 72: S3.I1&2 - Ergebnisse der Interviewerhebung, Wilcoxon-Test der Einzelzählungen der verschiedenen Teilbereiche, *N* = 10 240

Tabelle 73: S3.F1&2 - Fragestellungen des Blocks *Verständnisfragen* (VF) 246

Tabelle 74: S3.F1&2 - Fragestellungen des Blocks *Introspektion* (IS) 246

Tabelle 75: S3.F1&2 - Fragestellungen des Blocks *Einflussfaktoren* (EF) 247

Tabelle 76: S3.F1&2 - Fragestellungen des Blocks *Bewertung und Gewichtung* (BG) 248

Tabelle 77: S3.F1&2 - Abfolge der Frageblöcke im finalen Fragebogen 249

Tabelle 78: S3.F1&2 – Leitende Untersuchungsfragen der einzelnen Fragenblöcke 250

Tabelle 79: S3.F1&2 - Erweiterte Kategorien der *Introspektion* (VS-ET-IS) zur Auswertung von Frage VF1 251

Tabelle 80: S3.F1&2 – Kategorien zum *Verständnis von kognitiven Verzerrungen* (KV) zur Auswertung von Frage EF3 253

Tabelle 81: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis Selbstreflexion* (VS), *N* = 16 254

Tabelle 82: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis Kritisches Denken* (VK), *N* = 16 256

Tabelle 83: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis kognitive Verzerrungen* (KV), *N* = 16 257

Tabelle 84: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *selbstreflexive Aspekte* (BF-RA), *N* = 16 258

Tabelle 85: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *epistemologische Aspekte* (BF-EP), *N* = 16 259

Tabelle 86: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *naturwissenschaftliche Aspekte* (BF-NW), *N* = 16 260

Tabelle 87: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der offenen Fragestellungen, *N* = 16 261

Tabelle 88: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Introspektion* (IS), *N* = 16 262

Tabelle 89: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Blocks *Introspektion* (IS), *N* = 16 263

Tabelle 90: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Einflussfaktoren 2 (EF2), *N* = 16 263

Tabelle 91: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Items EF2, *N* = 16 264

Tabelle 92: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Einflussfaktoren 4* (EF4), personenbezogene Einflüsse, *N* = 16 264

Tabelle 93: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Einflussfaktoren 4* (EF4), kognitive Verzerrungen, *N* = 16 266

Tabelle 94: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der Items zu EF4, *N* = 16 267

Tabelle 95: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Bewertung und Gewichtung* (BG), *N* = 16 268

Tabelle 96: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Blocks *Bewertung und Gewichtung (BG*), *N* = 16 269

Tabelle 97: S3.F1&2 - Ergebnisse der Fragebogenerhebung, SRIS nach Silvia (2021), *N* = 16 269

Tabelle 98: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der SRIS, *N* = 16 271

Tabelle 99: S3.F1&2 - Ergebnisse der Fragebogenerhebung, CRT nach Frederick (2005), *N* = 14 271

Tabelle 100: S3.I1&2 – Interkoderreliabilität der Interviewerhebung 272

Tabelle 101: S3.F1&2 – Interkoderreliabilität der Fragebogenerhebung 273

Tabelle 102: S4.F1&2 – Settings des Blocks *Bewertungsaspekte* (ASP) 282

Tabelle 103: S4.F1&2 – Fragestellungen des Blocks *Bewertungsaspekte* (ASP) 282

Tabelle 104: S4.F1&2 - Abfolge der Frageblöcke im finalen Fragebogen 283

Tabelle 105: S4.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *epistemologische Faktoren* (EP-BF), *N* = 17 und *N* = 14 285

Tabelle 106: S4.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *naturwissenschaftliche Faktoren* (BF-NW), *N* = 17 und *N* = 14 286

Tabelle 107: S4.F1&2 - Ergebnisse der Fragebogenerhebung, Mann-Whitney-U-Test der offenen Fragen zum *Verständnis* (VS, VK, KV) und zu den *Bewertungsfaktoren* (BF), *N* = 17 und *N* = 14 287

Tabelle 108: S4.F1&2 - Ergebnisse der Fragebogenerhebung, Mann-Whitney-U-Test der geschlossenen Items, *N* = 17 und *N* = 14 288

Tabelle 109: S4.F1&2 – Ergebnisse der Fragebogenerhebung, CRT, *N* = 17 und *N* = 14 290

Tabelle 110: S4.F1&2 – Interkoderreliabilität der Fragebogenerhebung 290

Tabelle 111: S5.I1&2 - Ergebnisse der Interviewerhebung, *naturwissenschaftliche Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (NW), *N* = 7 294

Tabelle 112: S5.I1&2 - Ergebnisse der Interviewerhebung, *selbstreflexive Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (SR), *N* = 7 295

Tabelle 113: S5.I1&2 - Ergebnisse der Interviewerhebung, *epistemologische Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (EP), *N* = 7 298

Tabelle 114: S5.I1&2 - Ergebnisse der Interviewerhebung, *Verständnis Selbstreflexion* (SR-V), *N* = 7 299

Tabelle 115: S5.I1&2 - Ergebnisse der Interviewerhebung, *Verständnis Kritisches Denken* (KD-V), *N* = 7 300

Tabelle 116: S5.I1&2 - Ergebnisse der Interviewerhebung, *Selbsteinschätzung Fachwissen (NW-SF)* und *Fachwissenschaftliche Bewertung (NW-FB*), *N* = 7 301

Tabelle 117: S3.I1&2 - Ergebnisse der Interviewerhebung, *Fähigkeit zur Bewertung (FB)*, *N* = 7 302

Tabelle 118: S5.I1&2 - Ergebnisse der Interviewerhebung, Übersicht über die Einzelzählungen der verschiedenen Teilbereiche, *N* = 7 303

Tabelle 119: S5.I1&2 - Ergebnisse der Interviewerhebung, Wilcoxon-Test der Einzelzählungen der verschiedenen Teilbereiche, *N* = 7 303

Tabelle 120: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis Selbstreflexion* (VS), *N* = 12 309

Tabelle 121: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis Kritisches Denken* (VK), *N* = 12 311

Tabelle 122: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis kognitive Verzerrungen* (KV), ... 312

Tabelle 123: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *selbstreflexive Faktoren* (BF-SR), *N* = 12 313

Tabelle 124: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *epistemologische Faktoren* (BF-EP), *N* = 12 315

Tabelle 125: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *naturwissenschaftliche Faktoren* (BF-NW), *N* = 12 316

Tabelle 126: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der offenen Fragestellungen, *N* = 12 317

Tabelle 127: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Introspektion* (IS), *N* = 12 318

Tabelle 128: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Blocks *Introspektion* (IS), *N* = 12 319

Tabelle 129: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Einflussfaktoren 2 (EF2), *N* = 12 319

Tabelle 130: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Items EF2, *N* = 12 320

Tabelle 131: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Einflussfaktoren 4* (EF4), personenbezogene Einflüsse, *N* = 12 321

Tabelle 132: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Einflussfaktoren 4* (EF4), kognitive Verzerrungen, *N* = 12 322

Tabelle 133: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der Items zu EF4, *N* = 12 323

Tabelle 134: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Bewertung und Gewichtung* (BG), *N* = 12 324

Tabelle 135: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Blocks *Bewertung und Gewichtung (BG)*, *N* = 12 325

Tabelle 136: S5.F1&2 - Ergebnisse der Fragebogenerhebung, SRIS nach Silvia (2021), *N* = 12 326

Tabelle 137: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der SRIS, *N* = 12 327

Tabelle 138: S5.F1&2 - Ergebnisse der Fragebogenerhebung, CRT nach Frederick (2005), *N* = 16 328

Tabelle 139: S3.I1&2 – Interkoderreliabilität der Interviewerhebung 329

Tabelle 140: S3.F1&2 – Interkoderreliabilität der Fragebogenerhebung 330

Tabelle 141: Ergebnisse des Kruskal-Wallis-Tests für die Pre-Tests der Fragebogenerhebungen 337

Tabelle 142: Ergebnisse des Mann-Whitney-U-Tests für die Post-Tests der Fragebögen aus S3.F2 und S5.F2 340

Tabelle 143: Ergebnisse des Mann-Whitney-U-Tests für die Interviewerhebungen aus S3.I1&2 und S5.I1&2 342

Tabelle 144: Überblick über die untergeordneten Untersuchungsfragen des Projekts und deren Zuordnung zur Zielsetzung, Untersuchungsfrage 1 344

Tabelle 145: Überblick über die untergeordneten Untersuchungsfragen des Projekts und deren Zuordnung zur Zielsetzung, Untersuchungsfrage 2 344

Tabelle 146: Überblick über die untergeordneten Untersuchungsfragen des Projekts und deren Zuordnung zur Zielsetzung, Untersuchungsfrage 3 (Itemblöcke mit geschlossenen Fragestellungen wurden mit einem Sternchen * markiert) 345

Abkürzungsverzeichnis

ABM:	agenten-basierte Modellierung
ACT:	American College Test
CRT:	Cognitive Reflection Test
KI:	Künstliche Intelligenz
NOS:	Nature of Science
NOSI:	Nature of Scientific Inquiry
NOSIS:	Nature of Science-in-Society
NOSK:	Nature of Scientific Knowledge
SAT:	Scholastic Achievement Test
SRIS:	Self-Reflection and Insight Scale
SrS:	Systemisch-reflexive Stellungnahme
SSI:	Socioscientific Issues
SSR:	Socioscientific Reasoning
SuS:	Schülerinnen und Schüler
TPB:	Theory of Planned Behaviour
TPR:	Theory of Planned Reasoning
WHO:	World Health Organisation

Kurzzusammenfassung

Die zentrale Zielsetzung der naturwissenschaftlichen Grundbildung liegt nicht allein in der Vermittlung von Wissen begründet, sondern strebt die Befähigung der Lernenden an, naturwissenschaftliches Wissen auch in Handeln nutzbar zu machen. Diese Zielsetzung gewinnt insbesondere in der heutigen Zeit an Bedeutung, da naturwissenschaftliche Erkenntnisse zunehmend eine maßgebliche Rolle in gesellschaftlichen Entscheidungsprozessen einnehmen. Dennoch zeigt sich, dass vor allem bei komplexen gesellschaftlichen Problemstellungen mit naturwissenschaftlichem Bezug vorrangig affektive Faktoren, wie Emotionen, Werte oder soziale Normen, die Entscheidungsfindung entscheidend beeinflussen und zu Handlungen entgegen besseren Wissens führen können. Zusätzlich zu einer inhaltlichen Grundlage kommt damit dem Kritischen Denken eine zunehmende Relevanz zu. Obwohl es in der Bildung hervorgehoben und als Schlüsselkompetenz benannt wird, verbleibt Kritisches Denken jedoch ambig definiert, was eine gezielte Förderung erschwert. Vor diesem Hintergrund wird auf Grundlage verschiedener, auch interdisziplinärer, Perspektiven eine systematische Neudefinition Kritischen Denkens entwickelt und die Selbstreflexion im Denkprozess als vermittelnde Komponente zwischen affektiver und kognitiver Ebene hervorgehoben. Anschließend an diese Definition wird der Frage nachgegangen, wie sich Kritisches Denken nicht nur konzeptionell beschreiben, sondern auch empirisch erfassen lässt, um das Konstrukt didaktisch nutzbar zu machen. In diesem Rahmen werden fünf durchgeführte empirische Studien mit Chemielehramtsstudierenden der Universität zu Köln vorgestellt und ausgewertet. Die empirischen Befunde unterstützen die Annahme, dass Kritisches Denken durch die Analyse von Denk- und Bewertungsperspektiven in konkreten Bewertungsanlässen operationalisiert werden kann. Ferner zeigen die Ergebnisse, dass die Studierenden, unabhängig von ihrem Studienfortschritt, nur wenige unterschiedliche fachliche, epistemologische oder selbstreflexive Perspektiven in ihren Bewertungsprozess einbeziehen. Dies bestärkt die Hypothese, dass eine Erweiterung des inhaltlichen Wissens nicht automatisch mit einer Verbesserung multiperspektivischen Denkens und damit einer breiteren Bewertungsgrundlage einhergeht. Die Erprobung einer eigens dafür konzipierten Vermittlungseinheit zur expliziten Thematisierung Kritischen Denkens zeigt hingegen im Pre-Post-Vergleich signifikante Verbesserungen in den Bewertungsstrategien der Teilnehmenden. Damit wird die Notwendigkeit deutlich, Kritisches Denken explizit in Lehr-Lern-Prozesse zu integrieren, um nicht nur ein inhaltliches Verfügungswissen, sondern auch ein anschlussfähiges Orientierungswis-

sen zu fördern. Die Ergebnisse verdeutlichen, dass Kritisches Denken nicht nur theoretisch fassbar ist, sondern auch für die naturwissenschaftliche Vermittlung operationalisiert werden kann.

Dies eröffnet Anschlussmöglichkeiten für zukünftige Forschungen, um Kritisches Denken in tragfähige didaktische Konzepte zu überführen und es als festen Bestandteil naturwissenschaftlicher Bildung zu verankern.

Abstract

The central objective of scientific literacy is not just to impart knowledge, but to enable learners to apply scientific knowledge in their everyday lives. This objective is becoming an even more important goal in today's world, as scientific knowledge is playing an increasingly influential role in social decision-making processes. Nevertheless, it has been shown that, particularly in the case of complex socioscientific issues, affective factors, such as emotions, values or social norms, have a crucial influence on decision-making and can lead to actions contrary to better knowledge. In addition to a content-related basis, Critical Thinking is therefore becoming increasingly relevant. Although Critical Thinking is emphasized in education and named as a key competence, it is ambiguously defined, which makes targeted fostering difficult. As a response to this, this thesis proposes a systematic re-definition of Critical Thinking based on various perspectives, including interdisciplinary ones, and emphasizes self-reflection in the thinking process as a mediating component between the affective and cognitive levels. This definition is used as a basis for addressing the question of how Critical Thinking can not only be characterized conceptually but also be empirically evaluated in order to make the construct accessible for didactic use. In this context, five empirical studies were conducted and analyzed with chemistry teacher students at the University of Cologne. The findings support the assumption that Critical Thinking can be operationalized by analyzing thinking and assessment perspectives in concrete evaluation scenarios. Furthermore, the results indicate that, regardless of their academic progress, students only include a few different subject-specific, epistemological, or self-reflective perspectives in their assessment process. This supports the assumption that an increase in factual knowledge is not automatically accompanied by an improvement in multi-perspective thinking and thus a broader basis for assessment. On the other hand, the implementation of a course unit specifically designed to address Critical Thinking explicitly demonstrates significant improvements in participants' assessment strategies in a pre-post comparison. This clearly demonstrates the need to explicitly incorporate Critical Thinking into teaching and learning processes to promote not only content-related knowledge, but also orientational knowledge that can be applied. The results clearly show that Critical Thinking is not only conceptualizable in theory but can also be operationalized for science education. This paves the way for future research to transfer Critical Thinking into sustainable didactic concepts and anchor it as an integral part of science education.

Neuer Tag, neues Glück.

Gnothi seauton

1 Einleitung

> *Education should have two objects: first, to give definite knowledge - reading and writing, languages and mathematics, and so on; secondly, to create those mental habits which will enable people to acquire knowledge and form sound judgments for themselves. The first of these we may call information, the second intelligence.*
>
> \- Bertrand Russel (1922)

Die Naturwissenschaften, insbesondere die Chemie, prägen unser alltägliches Leben maßgeblich. Dabei trägt die Chemie nicht nur zu unserem wirtschaftlichen und ökologischen Fortschritt bei, indem sie die Grundlage vieler Professionen und Forschungsfelder bildet, sondern spiegelt als Kulturleistung auch unsere gesellschaftliche Entwicklung wider (Janich, 1992). Die Vermittlung und Thematisierung chemischer Inhalte trägt darüber hinaus zu einem naturwissenschaftlichen Grundverständnis bei, welches nicht nur der*m Einzelnen im Umgang mit modernen Technologien zugutekommen, sondern auch zu einer informierten gesellschaftlichen Teilhabe in partizipatorischen Prozessen beitragen soll. Diese partizipatorische Bedeutung der Naturwissenschaften im Allgemeinen und der Chemie im Besonderen wird auch in den aktuellen weiterentwickelten Bildungsstandards des Faches Chemie hervorgehoben: „Bildung in den Naturwissenschaften ermöglicht dem Individuum eine kenntnisreiche Teilhabe an gesellschaftlicher Kommunikation und Meinungsbildung über Potentiale und Risiken naturwissenschaftlicher Forschung und technische Entwicklungen“ (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024b, S. 5). Dennoch ist die Erweiterung und Übertragung einer im schulischen Kontext erworbenen Wissensgrundlage in die Lebensrealität mit Herausforderungen verbunden, die dem Ziel der naturwissenschaftlichen Grundbildung, Wissen in Handeln nutzbar zu machen, gegenüberstehen. Diese Herausforderungen werden im zweiten Kapitel dieser Arbeit aus verschiedenen Blickwinkeln beleuchtet.

Ein zentraler Aspekt in diesem Kontext ist das Fortschreiten der Digitalisierung, die durch die nahezu allgegenwärtige Verfügbarkeit und unregulierte Verbreitung von Informationen im Internet die Verantwortung für deren Bewertung in die Hände der Rezipierenden selbst legt (vgl. Kap. 2.1.3). Infolgedessen wird der Umgang mit digitalen Medien und Werkzeugen seitens der erweiterten Bildungsstandards als „integraler Bestandteil naturwissenschaftsbezogener Bildung“ (ebd., S. 5) betont und auch

die Bedeutung eines Verständnisses naturwissenschaftlicher Arbeits- und Denkweisen im Sinne von Nature of Science (NOS) benannt.

Dennoch zeigen Studien, dass bei der Auseinandersetzung mit Informationen und insbesondere in gesellschaftlich relevanten Problemstellungen mit naturwissenschaftlichem Bezug neben dem Wissen vor allem affektive Faktoren, wie persönliche Werte und Glaubenssätze (Cebesoy & Oztekin, 2018; Lee & Yang, 2019) oder die soziale Norm (Colombo et al., 2023; Dupont et al., 2024), einen entscheidenden Einfluss auf die Anwendung von Wissen nehmen (vgl. Kap. 2.1.6). Mögliche Erklärungsansätze liefern kognitionspsychologische Erkenntnisse, welche die Bedeutung der affektiven Ebene für das menschliche Denken betonen und zugleich ihren Einfluss auf die Genese und Rezeption wissenschaftlicher Erkenntnisse hervorheben (vgl. Kap. 2.3). Insbesondere in uneindeutigen und komplexen lebensweltlichen Entscheidungssituationen gewinnt damit die Förderung einer metakognitiven Denkpraxis zur Überwachung des eigenen Denkens an Bedeutung (Lee & Yang, 2019; Mestre & Cocking, 2002; Zeidler & Sadler, 2007). Damit verbunden birgt die Integration Kritischen Denkens in die naturwissenschaftliche Bildung das Potenzial, die eigenen Denkprozesse zu hinterfragen, kognitiven Verzerrungen entgegenzuwirken und die Qualität von Argumenten einzuschätzen, um Wissen bewusst einzusetzen und damit nutzbar zu machen (Barr & Ben-Shakhar, 2018).

Ursprünglich verfolgte diese Arbeit das Ziel, Kritisches Denken in der Ausbildung von Chemielehrkräften explizit zu fördern. Allerdings erwies sich der Begriff trotz seiner Hervorhebung als Schlüsselkompetenz für lebenslanges Lernen (Rat der Europäischen Union, 2018) sowie seiner expliziten und impliziten Benennung in nationalen und internationalen Bildungsdokumenten als mehrdeutig und konzeptuell unscharf (vgl. Kap. 2.2), was sowohl die Entwicklung von Förderansätzen als auch die empirische Erfassung ihrer Wirkung erschwerte. Aufgrund dessen verlagerte sich die Zielsetzung der Arbeit darauf, eine Grundlage in Form einer greifbaren und operationalisierbaren Definition Kritischen Denkens für die Naturwissenschaftsdidaktik bereitzustellen, um so einen Beitrag zu dessen Förderung zu leisten. Damit besteht die Arbeit aus zwei elementaren Teilen: Der konzeptionelle Teil widmet sich in Kapitel 3 der Definition Kritischen Denkens, indem bestehende Definitionen, sowohl aus der Naturwissenschaftsdidaktik als auch darüber hinaus, beleuchtet und Kernelemente herausgearbeitet werden. Um auch die affektive Ebene des Denkens zu integrieren, findet unter Hinzunahme interdisziplinärer kognitionspsychologischer Erkenntnisse darauf aufbauend eine Neudefinition Kritischen Denkens und eine systematische Übertragung auf die Naturwissenschaftsdidaktik statt.

Der empirische Teil der Arbeit knüpft in Kapitel 4 und 5 an die erarbeitete Definition an und adressiert die schrittweise Operationalisierung des Konstrukts in fünf iterativen Studien. Während das Vorgehen in den Vorstudien S1 und S2 (vgl. Kap. 5.1 und 5.2) in der frühen Phase der Entwicklung des Instruments noch stark explorativ ist und schwerpunktmäßig zur Ausschärfung der Definition beiträgt, erfolgt in den Studien S3 bis S5 (vgl. Kap. 5.3 bis 5.5) eine Überprüfung der Tragfähigkeit der Definition und des bis dahin entwickelten Instruments. Im Zuge dessen wird eine selbstkonzipierte Vermittlungseinheit zur expliziten Thematisierung Kritischen Denkens vorgestellt und das Instrumentarium daran erprobt. Zur Absicherung der Ergebnisse werden neben den intragruppalen Untersuchungen auch stichprobenübergreifende Vergleiche vorgenommen und die Ergebnisse in Kap. 5.6 vorgestellt. Das abschließende Fazit beantwortet die handlungsleitenden Forschungsfragen in Kapitel 6 und geht auf die Implikationen für die naturwissenschaftliche Bildung und die Chemiedidaktik ein, bevor ein Ausblick auf zukünftige Forschungsansätze gegeben wird.

Durch die Definition und Operationalisierung Kritischen Denkens sollen insbesondere für das Fach Chemie Perspektiven aufgezeigt werden, wie das sowohl im universitären als auch schulischen Kontext erworbene Anschlusswissen gezielt unterstützt und durch metakognitive Prozesse erweitert werden kann. Ziel ist es, in reflektierter Weise einen multiperspektivischen Möglichkeitsraum zu schaffen, in dem Wissen tatsächlich als Bewertungsgrundlage zur Anwendung kommen kann, und einen Übergang vom Inhalt zur Kompetenz, vom Verfügungs- zum Orientierungswissen und, in Anlehnung an Russels (1922) Zitat, von der Information zur Intelligenz zu fördern.

2 Theoretische Grundlegungen

Kritisches Denken stellt gerade in unserer heutigen Zeit ein immanent wichtiges metakognitives Konstrukt dar, welches sich, obgleich seiner trivialen Verwendung im Alltagssprachgebrauch, hochkomplex darstellt. Angesichts dessen ist eine Definition der Begrifflichkeit Kritischen Denkens unumgänglich, bevor eine Operationalisierung erfolgen kann. Um aber zunächst die Relevanz Kritischen Denkens für die Chemiedidaktik aufzuzeigen, ist es notwendig, nicht nur fachlich chemische oder chemiedidaktische Inhalte zu betrachten, sondern an die Wurzel dessen zu gehen, was die Fachdidaktik zu adressieren sucht: das Denken selbst. Hierzu werden über die Grenzen der Fachdidaktik hinweg auch Inhalte aus den Bereichen der Kognitions- und Verhaltenspsychologie herangezogen, um einen systemischen Blick auf das Konstrukt des Kritischen Denkens zu erhalten, bevor dieses auf konkrete chemische Inhalte angewendet wird.

2.1 Die Überführung von Wissen in Handeln

Die Zielsetzung von Bildung hat sich in den vergangenen Jahrzehnten grundlegend verändert. Von einer anfänglichen Inhaltsorientierung ging der Schritt über die Lernzielorientierung hin zu einer Förderung von Kompetenzen, welchen nicht nur eine Inhalts-, sondern auch eine Handlungskomponente innewohnt. Von dieser Kompetenzorientierung und dem damit einhergehenden Ziel, Wissen in Handeln zu überführen, ausgehend, fanden auch erweiternde Konzepte im Lehr-/Lernkontext immer mehr Berücksichtigung, beispielsweise die Thematisierung von Nature of Science (NOS) und eine erhöhte Sensibilisierung für digitale Kompetenzen. Dennoch zeigen sich gerade in lebensweltlichen Kontexten Herausforderungen, das in den Naturwissenschaften und speziell im Fach Chemie erlernte Wissen in gesellschaftliche, ökologische, ökonomische oder persönliche Entscheidungen miteinzubeziehen. Im Folgenden wird ein Überblick über die aktuellen Bemühungen und Herausforderungen gegeben, die dem Bildungsziel, Wissen effektiv nutzbar zu machen, gegenüberstehen.

2.1.1 Ziele naturwissenschaftlicher Grundbildung

Die Vermittlung von Wissen erfolgt nicht als Selbstzweck, sondern strebt danach, dass sich das erworbene Wissen in der Anwendung bewährt und dadurch seinen Nutzen entfaltet. Schon Comenius schrieb im neunten seiner Grundsätze zu leichtem Lehren und Lernen im 17. Kapitel der *Didactica Magna*:

> Die Natur erzeugt nichts, dessen Nutzen nicht bald offenbar wird. […] Deshalb wird man dem Schüler seine Arbeit erleichtern, wenn man ihm zeigt, welches der Nutzen von dem, was man ihn lehrt, im täglichen Leben ist. Daran muss man auf allen Gebieten festhalten, in der Grammatik, Dialektik, Arithmetik, Geometrie, Physik usw. (Comenius & Overhoff, 2022, S. 201 f.; Pappenheim, 1892, S. 150).

Auch Dewey (2011, nach dem Original von 1916) begriff Bildung in seinem humanistischen Verständnis nicht als reine Aneignung von Wissen und Fähigkeiten, sondern vor allem als Beitrag zur Entwicklung der*s Einzelnen als Teil der Gesellschaft. Für ihn war Bildung ein lebenslanger Prozess mit Beitrag zur Persönlichkeitsentwicklung und dem Ziel einer verantwortungsbewussten Mündigkeit zur Partizipation an gesellschaftlichen Prozessen. Bis zur Bildungsreform 1957, welche durch den Sputnik-Schock, die Reaktion des Westens auf den erfolgreichen Satellitenstart der damaligen Sowjetunion, angestoßen wurde, waren die Lehrpläne in Deutschland jedoch schwerpunktmäßig inhaltsorientiert ausgerichtet (Smidt, 2017). Erst im Zuge dieser wurden die gesellschaftlichen Implikationen der Bildung vermehrt ins Auge gefasst und auch der persönliche Entwicklungsstand der Lernenden berücksichtigt (Reiners, 2022). Dies schlug sich in der Abkehr von der Inhaltsorientierung der Lehrpläne und Zuwendung zur Lernzielorientierung in Lernplänen nieder (ebd.). Ein Lernziel lässt sich dabei nach Meyer (1977) definieren als „sprachlich artikulierte Vorstellung über die durch Unterricht (oder andere Lehrveranstaltungen) zu bewirkende Verhaltensänderung eines Lernenden“ (ebd., S. 21). Zusätzlich zur Inhaltsebene enthält das Lernziel damit eine Verhaltenskomponente, welche konkretisiert, wie mit dem Wissen umgegangen werden soll (ebd.).

Ausgelöst durch die unterdurchschnittlichen PISA-Ergebnisse der 2000er Jahre (Prenzel et al., 2001; Reiners, 2022; Smidt, 2017) wurde die Forderung nach einer weiteren Bildungsreform laut, welche den Schritt von der Lernziel- hin zur Kompetenzorientierung anstieß, die 2004 Einzug in die deutschen Bildungsstandards hielt (Kultusministerkonferenz, 2005). Der Kompetenzbegriff spiegelt dabei das Ziel wider, Wissen in Handeln nutzbar zu machen, da in ihm sowohl eine Inhaltsebene als auch eine konkrete Handlungsebene impliziert ist. Nach der in den Bildungsstandards (Kultusministerkonferenz, 2005; Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2020) und den PISA-Studien (Prenzel et al., 2001; Rychen & Salganik, 2001) verwendeten Definition nach Weinert (2001) sind Kompetenzen definiert als

> […] die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können (ebd., S.27).

Diese Definition legt den Schwerpunkt auf erworbene Fähigkeiten und deren Nutzung in der Lebensrealität und versteht Kompetenz als Disposition zur Befähigung der Problemlösung in konkreten Anforderungssituationen (Faulstich-Christ et al., 2010; Klieme et al., 2003). Auf diesem Verständnis fußt auch das Kompetenzmodell der aktuellen Bildungsstandards (Faulstich-Christ et al., 2010; Reiners, 2022; Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2020) mit dem Ziel der Befähigung der Lernenden zur gesellschaftlichen Teilhabe und Orientierung in der Lebenswelt. Dabei spiegelt eine Kompetenz mehr wider als nur Wissen und kognitive Fähigkeiten, sondern bedarf auch des Einsatzes von psychosozialen Ressourcen zur Bewältigung komplexer Aufgaben (Rychen & Salganik, 2001).

In den weiterentwickelten Bildungsstandards im Fach Chemie wird dezidiert der angestrebte Beitrag der Naturwissenschaften im Allgemeinen und der Chemie im Besonderen zur Bildung hervorgehoben und detailliert aufgeführt (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024b). Die Naturwissenschaften sollen ein tiefes Verständnis für die zunehmend technologisierte Welt vermitteln, Lernende zur aktiven Teilnahme am gesellschaftlichen Diskurs befähigen und als integraler Bestandteil der kulturellen Identität verstanden werden. Die Vermittlung des Faches Chemie strebt insbesondere an, den Lernenden die Fähigkeit zu vermitteln, ihre Lebenswelt sowohl auf makroskopischer als auch auf submikroskopischer Ebene zu begreifen, und sie zu befähigen, fundierte Entscheidungen zu treffen, die in individuellen, gesellschaftlichen, ökonomischen und ökologischen Kontexten von Bedeutung sind. Ebenso wird die Förderung eines Verständnisses der naturwissenschaftlichen Methoden der Erkenntnisgewinnung angestrebt, um die Reflexion über die Potenziale und Grenzen naturwissenschaftlicher Arbeits- und Denkweisen zu ermöglichen. Ziel der chemiespezifischen Bildung ist es, das Bewusstsein für die nachhaltige Nutzung von Ressourcen zu schärfen und die Verantwortung im Umgang mit Stoffen und Technologien zu fördern. Zudem wird der interdisziplinäre Charakter der Chemie betont, der Verbindungen zu anderen naturwissenschaftlichen Disziplinen sowie gesellschaftsrelevanten Themen

schafft und so die Grundlage für eine ganzheitliche naturwissenschaftliche Bildung legt (ebd.).

Zur Erreichung dieser Ziele umfassen die aktuellen Bildungsstandards im Fach Chemie vier Kompetenzbereiche:

- Die Sachkompetenz umfasst das Wissen über naturwissenschaftliche Begrifflichkeiten und Konzepte, Theorien und Verfahrensweisen sowie die Fähigkeit, diese zu beschreiben und auf alltagsbezogene Sachverhalte zu übertragen.

- Die Erkenntnisgewinnungskompetenz beinhaltet das Verständnis grundlegender naturwissenschaftlicher Arbeits- und Denkweisen sowie die Fähigkeit zur Reflexion der Möglichkeiten und Grenzen ebenjener.

- Die Kommunikationskompetenz schließt die Fachsprache und fachtypische Darstellungen ein sowie die Fähigkeit, fachliche Informationen zu erschließen und adäquat zu verwenden.

- Die Bewertungskompetenz enthält zu guter Letzt das Wissen um fachliche und überfachliche Perspektiven zur Bewertung in Verbindung mit der Fähigkeit, Urteile kriteriengeleitet fällen und auch ethische Konsequenzen abwägen zu können (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024b, 2020).

Vor allem der Bereich der Sachkompetenz hat sich seit 2005 auch in seiner Bezeichnung, vom Kompetenzbereich *Fachwissen* hin zur *Sachkompetenz,* dahingehend weiterentwickelt, dass der Schwerpunkt vom Erwerb des Wissens innerhalb der chemischen Domäne hin zu einem ganzheitlicheren Verständnis des Wissens und dessen Anwendung in realen und lebensweltlichen Kontexten verlagert wurde. Dies unterstreicht das Bestreben, nicht nur rein inhaltlich zu fördern, sondern auch eine aus dem Unterrichtsfach Chemie erwachsende Handlungsfähigkeit anzustreben.

Diese Handlungsorientierung findet sich ebenfalls in der OECD-Definition der naturwissenschaftlichen Grundbildung (Scientific Literacy) wider, welcher neben anderen Definitionen aufgrund ihrer stetigen Berücksichtigung innerhalb der PISA-Studien eine vorrangige Relevanz zukommt (Osborne, 2023):

> Naturwissenschaftliche Grundbildung (Scientific Literacy) ist die Fähigkeit, naturwissenschaftliches Wissen anzuwenden, naturwissenschaftliche Fragen zu erkennen und aus Belegen Schlussfolgerungen zu ziehen, um Entscheidungen zu verstehen und zu treffen, die die natürliche Welt und die durch menschliches Handeln an ihr vorgenommenen Veränderungen betreffen (Organisation for Economic Co-operation and Development, 1999, S. 60; nach der Übersetzung von Prenzel et al., 2001, S. 198).

Die naturwissenschaftliche Grundbildung[1] stellt dabei das in der Naturwissenschaftsdidaktik zentrale Bildungsziel dar (Gräber & Nentwig, 2002) und wird hier nicht als Fachwissen beschrieben, sondern als die Fähigkeit, dieses zur Anwendung zu bringen und zur Orientierung nutzen zu können. Im Zuge der Allgemeinbildung soll damit kein aus den Fachdisziplinen zusammengesetztes Gesamtwissen entstehen, sondern eine „Transzendierung der Fachlichkeit" (Gräber, 2002, S. 7) sowie allgemeine Kompetenzen angestrebt werden. Dabei werden das sichere Beherrschen kultureller Basiswerkzeuge, wie Sprache oder mathematische Symbole, die Entwicklung von Orientierungswissen in zentralen Wissensdomänen, sozial-kognitive und soziale Kompetenzen sowie metakognitive Kompetenzen besonders hervorgehoben (BLK, 1997; Gräber, 2002).

Neben der OECD-Definition gibt es noch weitere Definitionsansätze, wie den von Schaefer (2002), welcher Scientific Literacy ebenfalls als integralen Bestandteil der Allgemeinbildung sieht und damit als „Lebenskompetenz" (Gräber, 2002, S. 10; Gräber & Nentwig, 2002, S. 14; Schaefer, 2002, S. 102) bezeichnet. Die Sachkompetenz und ihr Erwerb bilden dabei die Voraussetzung für weitere Kompetenzen, wie die Sozial-, ethische oder epistemologische Kompetenz. Diese soll kooperativ fächerübergreifend verfolgt werden. Er beschreibt zwei Stufen der Scientific Literacy:

1. Erwerb von Grundwissen und Fähigkeiten als statische Aspekte des Lernens,
2. Aktivierung und Anwendung von Wissen und Fähigkeiten als dynamischer Aspekt des Lernens.

Dubs (2002) indes fasst Scientific Literacy als die Fähigkeit auf, naturwissenschaftlich-technische Konfliktthemen zu verstehen und darüber begründet argumentieren zu können, was insbesondere das Erkennen von Interessenkonflikten und Wertvorstellungen miteinschließt. Ziel ist auch hier, Problemstellungen in der Lebensrealität

[1] Im Weiteren wird aufgrund der schwerpunktmäßigen Verwendung des Begriffs in der Forschungsliteratur von *Scientific Literacy* gesprochen, was dem deutschen Begriff der naturwissenschaftlichen Grundbildung entspricht.

wahrnehmen und analysieren sowie eigene Entscheidungen in wertbezogenen Fragestellungen treffen zu können, wobei nicht allen Problemen rationale Lösungen innewohnen (Gräber, 2002).

Bybee (1997, 2002) wiederum unterscheidet vier Dimensionen der Scientific Literacy, welche hierarchisch strukturiert sind. Die *nominale* Scientific Literacy beschreibt ein naives Verständnis der Naturwissenschaften, jedoch mit der Fähigkeit, naturwissenschaftliche Fragen ausmachen zu können, wohingegen die *funktionale* Scientific Literacy eine adäquate Verwendung naturwissenschaftlicher Begrifflichkeiten impliziert. Die *konzeptionelle* oder auch *prozedurale* Scientific Literacy spiegelt ein allgemeines Verständnis der Konzepte und Prinzipien der Naturwissenschaft wider und in letzter Instanz kann unter der *multidimensionalen* Scientific Literacy ein Verständnis der Besonderheiten der Naturwissenschaften und deren Verflechtungen mit sozialen Kontexten gefasst werden. Diese unterschiedlichen Dimensionen verdeutlichen, dass Scientific Literacy nicht nur auf das reine Wissen begrenzt ist, sondern auch das Verstehen der Einbettung von Wissenschaft in gesellschaftliche Diskurse erfordert (S. Müller, 2021).

Aus diesem Spannungsgefüge der Definitionen für die naturwissenschaftliche Grundbildung hervorgehend formulieren Bybee und Roberts (2014) zwei Visionen: die grundlagen- und inhaltsorientierte *Vision I*, deren Zielsetzung das Ausbilden eines Verständnisfundaments zum späteren Weiterlernen der Naturwissenschaft ist (interner Blick auf die Naturwissenschaften), und die anwendungsbezogene und verknüpfte *Vision II,* welche die Befähigung zur Anwendung in komplexen Bewertungssituationen auch in Verbindung mit der Gesellschaft anstrebt (externe Zusammenhänge der Naturwissenschaften). Osborne (2023) benennt die *Vision II* radikaler als Ziel, wissenschaftliches Wissen in Handeln umsetzen zu können (Bencze et al., 2020; Sjöström & Eilks, 2020) und ergänzt noch um eine *Vision III*, die naturwissenschaftliche Grundbildung als kollektives Gut versteht, welches in seinem Umfang nicht von einer einzelnen Person vollständig erreicht werden kann (National Academies of Sciences, Engineering, and Medicine, 2016). Diese Visionen veranschaulichen, dass Scientific Literacy nicht nur individuelle Kompetenzen umfasst, sondern auch die Fähigkeit, die Naturwissenschaften in sozialen Kontexten zu greifen und einordnen zu können.

Zusammenfassend definieren Bybee und Roberts (2014) Scientific Literacy wie folgt:

> Scientific literacy is the ability to engage with sciencerelated issues and with the ideas of science as a reflective citizen. A scientifically literate person, therefore, is willing to engage with science-related issues in reasoned discourse about science and technology, which requires the competencies to:
>
> 1. explain phenomena scientifically: recognize, offer and evaluate explanations for a range of natural and technological phenomena
>
> 2. understand scientific enquiry: recognize, describe and evaluate the design, practices and conduct of scientific enquiry
>
> 3. interpret scientific evidence: analyze and evaluate scientific information, claims, and arguments in a variety of representations and draw appropriate conclusions (ebd., S. 552).

In dieser Definition werden sowohl die Bereitschaft als auch die Fähigkeiten betont, die nötig sind, um naturwissenschaftliches Wissen zur Anwendung zu bringen, was sowohl der *Vision I* als auch der *Vision II* Rechnung trägt. Letztere erfuhr jedoch in den 2000, 2006 und 2015 durchgeführten PISA-Studien immer weniger Beachtung (Bybee & Roberts, 2014; L. Müller, 2021). Dies ist bedenklich, da Lernende im Alltag den Naturwissenschaften nicht in ihrer reinen Form, sondern in persönlichen, gesellschaftlichen und globalen Kontexten begegnen (Bybee & Roberts, 2014) und daher gerade dieser *Vision II* für die Lebensrealität der Lernenden ein wesentlicher Wert zukommt. Shamos (1995, 2002) beschreibt eine schwerpunktmäßig inhaltliche Scientific Literacy gar als unerreichbar und spricht sich für einen verstärkten Fokus auf ein Verständnis der Methoden und Prozesse der naturwissenschaftlichen Erkenntnisgewinnung aus, um den Lernenden ein Verständnis für das Wesen der Naturwissenschaften zu vermitteln.

Auf Grundlage der allgemeinen Bildungszielsetzung und der speziellen Ausrichtung der Naturwissenschafts- und Chemiedidaktik, Wissen nicht nur zu vermitteln, sondern es in Handeln umsetzbar zu machen, sowie unter Berücksichtigung der Definitionen von naturwissenschaftlicher Grundbildung und des Kompetenzbegriffs lassen sich zwei Arten von Wissen unterscheiden: Das *Verfügungswissen* (Sachwissen), welches dazu befähigt, Dinge zu tun, und das *Orientierungswissen* (Wissen um Handlungsmaßstäbe), welches ermöglicht, komplexe und moralische Entscheidungen zu treffen und sich in der Lebenswelt zu orientieren (Mohr, 1989; Reiners, 2022). Dabei

kann argumentiert werden, dass die *Vision I* der Scientific Literacy schwerpunktmäßig dem *Verfügungswissen* entspricht und die *Vision II* das *Orientierungswissen* anstrebt. Dabei ist hervorzuheben, dass beide Wissensarten, obwohl sie aus unterschiedlichen Quellen hervorgehen, gleichermaßen ihre Berechtigung haben. Gleich den beiden Visionen von Scientific Literacy bedarf es beider Ansprüche: dem grundlegenden Verständnis und der Fähigkeit, dieses auch nutzen zu können.

2.1.2 Die Bedeutung von Nature of Science

Im Rahmen der Bestrebungen, Kompetenzen zu fördern, Orientierungswissen zu vermitteln und Lernende durch die Vermittlung einer naturwissenschaftlichen Grundbildung zur gesellschaftlichen Teilhabe zu befähigen, gewinnt auch die Vermittlung eines Verständnisses der Charakteristika der Naturwissenschaften an Relevanz, da nur so die Bedeutungen, Grenzen, Möglichkeiten und Potenziale ebenjener eingeschätzt werden können (S. Müller, 2021; Shamos, 1995, 2002). Die Reflexion dieser naturwissenschaftlichen Arbeits- und Denkweisen und der implizierten Einstellungen und Werte wird unter der Begrifflichkeit *Nature of Science* (NOS) gefasst und beschreibt eine Art Metawissen der Naturwissenschaften (Lederman, 1992; Reiners, 2022). Eine genaue Definition lässt sich aufgrund der ihr innewohnenden und in ihrer Natur begründeten Vorläufigkeit schwer fassen (ebd.), jedoch lassen sich Charakteristika formulieren, die als Kernelemente herausgearbeitet werden können. Lederman (1992) formuliert sieben Aspekte, darunter die Vorläufigkeit, Empirie, Subjektivität, Kreativität und die Einbettung in soziale Kontexte. Ferner benennt er die Notwendigkeit, zwischen Beobachtung und Schlussfolgerung sowie Theorie und Gesetzen zu unterscheiden (Lederman, 1992; Reiners, 2022). Aufgrund der Diskussion, dass dieser Ansatz lediglich die naturwissenschaftlichen Erkenntnisse beschreibt, aber nicht den Weg der Erkenntnisgewinnung abdeckt, unterscheidet Lederman (2018) zwischen *Nature of Scientific Inquiry* (NOSI), welche die Prozesse und Methoden umfasst, durch die wissenschaftliches Wissen generiert wird, und *Nature of Scientific Knowledge* (NOSK), welche sich auf die Charakteristika naturwissenschaftlichen Wissens bezieht (Lederman, 2018; Lederman & Lederman, 2019). Die Lederman Seven wären damit schwerpunktmäßig NOSK zuzuordnen. Auch wenn neben NOS die Termini von NOSK und NOSI in der Literatur auftauchen, wird hier weiterhin nur die Bezeichnung NOS verwendet, da dies die anderen Bereiche mit umfasst und bei der Definition von NOSK und NOSI keine haltbare Einigung herrscht. Im Rahmen der Diskussion um die verschiedenen Charakteristika von NOS ist aber sowohl das Berücksichtigen der Charakteristika der wissenschaftlichen Erkenntnisse als auch der Prozesse der Erkenntnisgewinnung von Relevanz, um die Einbettung

naturwissenschaftlichen Wissens in gesellschaftliche und kulturelle Kontexte einordnen zu können.

McComas und Olson (1998) kommen bei der Charakterisierung von NOS zu vergleichbaren Ergebnissen wie Lederman und stimmen vor allem im Hinblick auf die Vorläufigkeit, das empirische Fundament, die Kreativität sowie soziokulturelle Einbettung mit Lederman überein. Die Delphi-Studie von Osborne et al. (2003) unterstreicht diese Charakteristika und erweitert diese um die naturwissenschaftliche Methodik sowie ihre Vielfältigkeit, stellt aber auch die Verknüpfung zwischen Naturwissenschaft und Technik heraus. Eine Übersicht über die Charakteristika und ihre Übereinstimmungen ist in Tabelle 1 zu finden (Albrecht, 2021; vgl. Neumann & Kremer, 2013).

Tabelle 1: Übersicht über die Kern-Elemente von NOS. Gegenüberstellung von Lederman, McComas & Olson und Osborne et al. (Albrecht, 2021; vgl. Neumann & Kremer, 2013)

Lederman Seven (Lederman, 1992; Reiners, 2022)	McComas & Olson (McComas & Olson, 1998)	Osborne et al. (Osborne et al., 2003)
Naturwissenschaftliche Erkenntnisse sind vorläufig	Naturwissenschaftliche Erkenntnisse sind vorläufig	Wissenschaft und Gewissheit (Vorläufigkeit und Beständigkeit in der Naturwissenschaft)
Naturwissenschaftliche Erkenntnisse sind empirisch belegt	Naturwissenschaftliche Erkenntnisse sind empirisch belegt	Analyse und Dateninterpretation (Rolle der Interpretation von Daten und ihrer Subjektivität sowie Ablauf der Bildung von Theorien)
Naturwissenschaftliche Erkenntnisse sind subjektiv		
Naturwissenschaftliche Erkenntnisse erfordern eine Unterscheidung in Beobachtung und Schlussfolgerung	Naturwissenschaftliche Erkenntnisse basieren auf Reproduzierbarkeit und ehrlicher Informationsdarlegung	Naturwissenschaftliche Methodik und kritisches Testen (Verwendung und Grenzen experimenteller Herangehensweisen)
Naturwissenschaftliche Erkenntnisse erfordern eine Unterscheidung zwischen Theorien und Gesetzen		Vielfältigkeit des naturwissenschaftlichen Ansatzes (Wissen um eine Bandbreite von Methoden und dass es nicht die eine naturwissenschaftliche Methode gibt)
	Naturwissenschaften versuchen Phänomene zu erklären	Hypothesen und Vorhersagen zu natürlichen Phänomenen
		Zyklischer Prozess der Fragestellung in den Naturwissenschaften
Naturwissenschaftliche Erkenntnisse schließen menschliche Vorstellungskraft und Kreativität mit ein	Naturwissenschaftler sind kreativ	Kreativität
Naturwissenschaftliche Erkenntnisse sind in soziokulturelle Kontexte eingebettet	Naturwissenschaften sind Teil einer gesellschaftlichen Tradition	Kooperation und Kollaboration zur Entwicklung naturwissenschaftlicher Erkenntnisse

Naturwissenschaftliche Ideen werden von gesellschaftlichen sowie historischen Kontexten beeinflusst	Historische Entwicklung naturwissenschaftlichen Wissens
Naturwissenschaften spielen eine wichtige Rolle in der Technologie	Naturwissenschaft und Technik (Verbindung und Abhängigkeit voneinander)
Naturwissenschaften haben eine globale Bedeutung	
Veränderungen in den Naturwissenschaften treten schrittweise ein	
Neues naturwissenschaftliches Wissen muss klar und offen kommuniziert werden	

Ein Verständnis von NOS fördert nicht nur ein Verständnis naturwissenschaftlicher Arbeits- und Denkweisen, sondern will auch ein Bewusstsein für die Art und Weise schaffen, wie Wissenschaft in gesellschaftliche und kulturelle Diskurse eingebettet ist. Damit leistet ein NOS-Verständnis einen Beitrag zur informierten Mündigkeit und Partizipation an der Demokratie (S. Müller, 2021). McComas et al. (2020) nennen dabei das *Human Element of Science* (ebd.) als eines der wesentlichen NOS-Cluster, welches auch in den Next Generation Science Standards der USA Erwähnung findet (Achieve, Inc., 2013), und stellen so die Bedeutung des Menschen und seiner Eingebundenheit bei der Genese von Wissen heraus (Abbildung 1).

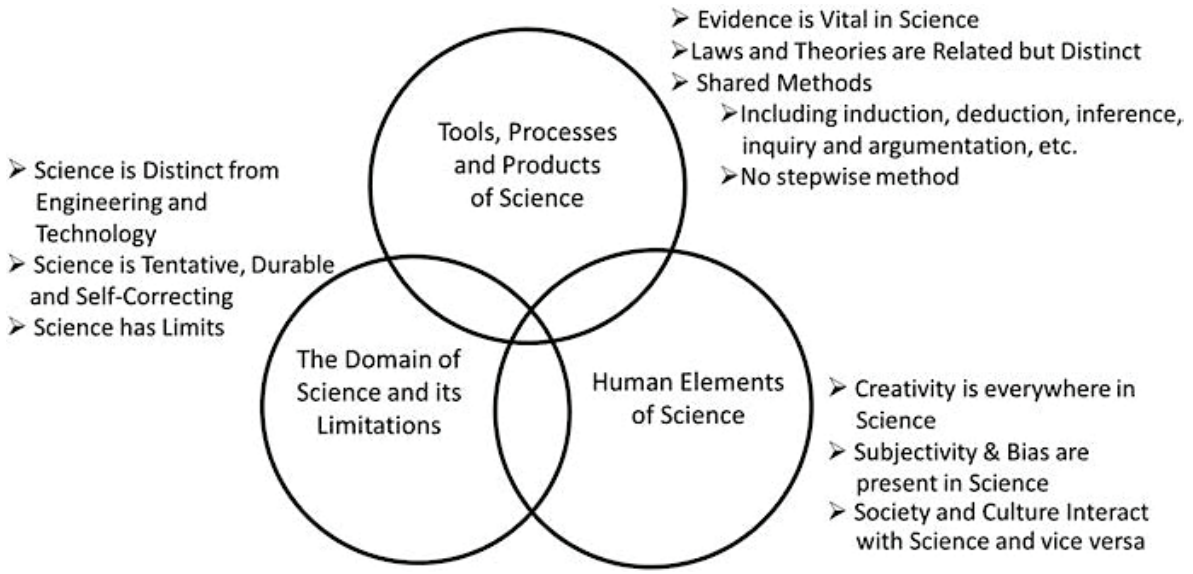


Abbildung 1: The nine major sub-elements or key NOS aspects (McComas, 2020, S. 40)

Driver et al. (1996) begründen die Vermittlung von NOS basierend auf Thomas und Durant (1987) mit fünf Argumenten:

- Das *ökonomische Argument* stellt die wirtschaftliche Notwendigkeit heraus, qualifizierte Naturwissenschaftler für eine Weiterentwicklung der Industrie auszubilden.
- Das *utilitaristische Argument* verdeutlicht die Bedeutung eines naturwissenschaftlichen Grundwissens, um sich in der heutigen technologisierten Welt zurechtfinden zu können.
- Das *demokratische Argument* hebt die Relevanz gesellschaftlicher Mündigkeit durch die Vermittlung eines Verständnisses der Naturwissenschaften hervor, da gesellschaftsrelevante Entscheidungen immer mehr naturwissenschaftliche und technische Fragestellungen berühren.
- Das *kulturelle Argument* sieht die Naturwissenschaften als schätzenswerte kulturelle Meilensteine an.
- Das *moralische Argument* führt die Wichtigkeit eines naturwissenschaftlichen Verständnisses für die Berücksichtigung von Moral und Werten an, da die Naturwissenschaften relevante ethische Fragestellungen aufwerfen.

All diese Argumente stützen die Implementierung von NOS in die Vermittlung der Naturwissenschaften. Besonders hervorzuheben sind jedoch das *demokratische* und das *moralische* Argument, welche dem Ziel der naturwissenschaftlichen Grundbildung, Lernende zu Mündigkeit zu erziehen, besonders entsprechen. Gesellschaftliche Fragestellungen mit einem naturwissenschaftlichen Bezug werden dabei mit dem wissenschaftlichen Fortschreiten zunehmend komplexer und auch der digitale Wandel erfordert eine vermehrte Berücksichtigung und Anpassung bestehender Konzepte und didaktischer Ansätze, was sich auch in der Anpassung von Bildungsstandards und Leitlinien niederschlägt (Sekretariat der Kultusministerkonferenz, 2016; Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024a, 2024b).

Eine solche Erweiterung stellt die Ergänzung der bestehenden NOS-Konzepte um die Wissenschaftskommunikation in der Gesellschaft dar, welche Höttecke und Allchin (2020) als NOSIS (*Nature of Science-in-Society*) bezeichnen. Während früher den klassischen Medien, wie Zeitung, Radio oder Fernsehen, durch die begrenzte Anzahl an Kanälen eine Art Vermittlerrolle zufiel und Wissenschaftskommunikatoren Inhalte für die breite Masse aufarbeiten konnten, bietet heute der fast allgegenwärtige Zugang zum Internet nahezu unbegrenzte Publikationsmöglichkeiten (ebd.). Infolgedessen sind Informationen nicht mehr auf Moderation oder Prüfung angewiesen, sondern können per Tastendruck direkt im Internet veröffentlicht werden, wodurch sie einem großen Publikum zugänglich gemacht werden.

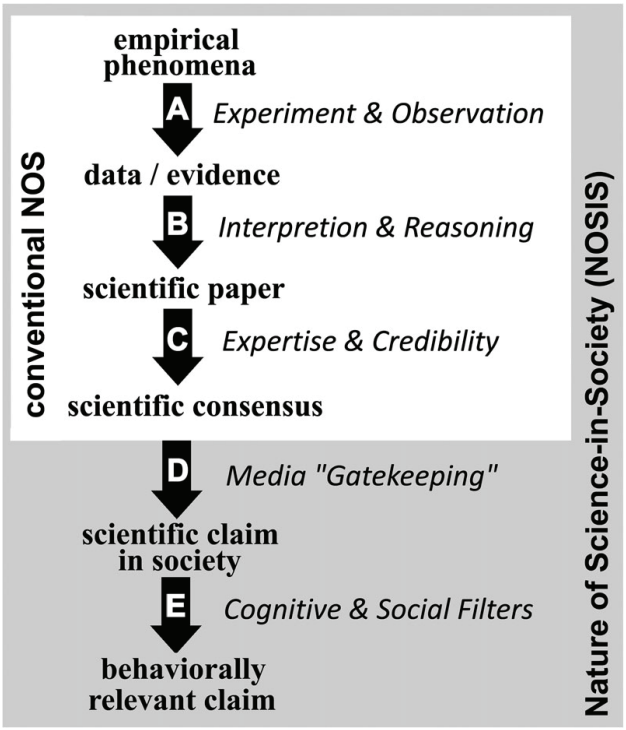


Abbildung 2: Nature of Science-in-Society (NOSIS) (Höttecke & Allchin, 2020, S. 5)

Des Weiteren sind bei der Rezeption von Informationen die kognitiven und sozialen Filter[2] von großer Bedeutung, da diese im Wissenschaftskommunikationsprozess nur schwerlich beeinflusst werden können (Abbildung 2).

Angesichts der Möglichkeiten und Herausforderungen der neuen Medien obliegt es den Rezipierenden, Informationen auf Verlässlichkeit zu prüfen und die Herkunft dieser nachzuvollziehen (ebd.). Eine fundierte Einordnung kann allerdings nur dann erfolgen, wenn ein Verständnis der Charakteristika der Naturwissenschaften vorhanden ist, die Möglichkeiten und Grenzen der Genese reflektiert werden können und ein Bewusstsein für die kognitiven und sozialen Filter vorliegt.

Die technologischen, naturwissenschaftlichen und gesellschaftlichen Entwicklungen unterstreichen die Herausforderungen, die mit der Anwendung von Wissen einhergehen, und zeigen in ihrem lebensweltlichen Bezug die Komplexität außerhalb des Klassenzimmers durch die Uneindeutigkeiten von Bewertungssituationen auf. Multifaktorielle Probleme bieten selten klare Lösungen und Einteilungen in richtig oder falsch. Daher stellt sich angesichts der Herausforderungen durch die Medien, den Einfluss von Fehlinformationen auf Politik und Gesellschaft sowie die anhaltende

[2] Auf diese wird im späteren Verlauf rekurriert.

Beliebtheit pseudowissenschaftlicher Ansätze die Frage, ob das Ziel, Wissen in Handeln umzusetzen, durch die aktuellen Bemühungen ausreichend erreicht werden kann.

2.1.3 Die Herausforderung durch die Medien

Die SARS-CoV-2-Epidemie zeigte deutlich, dass die modernen Medien, wie Social-Media-Kanäle, Podcasts oder Streams, nicht den klassischen Regeln der Informationsverarbeitung unterliegen. In den sozialen Medien kam es zu zahlreichen Verharmlosungen, unbegründeten und undifferenzierten Kritiken sowie negativen Bewertungen von Entscheidungen. Überdies bildeten sich Gegenbewegungen auf Basis von Fehlinformationen, die teilweise bis hin zu Hetzaufrufen gegen führende Wissenschaftler*innen führten (Hanfeld, 2020). Dr. Tedros Adhanom Ghebreyesus, zu diesem Zeitpunkt Generaldirektor der WHO (World Health Organisation), benannte die Situation am 15.02.2020 auf der Münchner Sicherheitskonferenz mit den Worten: „But we´re not just fighting an epidemic; we´re fighting an infodemic“ (Ghebreyesus, 2020, S. 4).

Während in klassischen Medien Wissenschaftsjournalist*innen als Vermittler*innen und Aufbereiter*innen naturwissenschaftlicher Informationen auftreten, führen die direkten Kommunikationswege moderner Medien zu einer breiteren und unkontrollierteren Verbreitung von Inhalten. Der direkte Zugang zu Informationen über das Internet bietet zwar Chancen, wie die ständige Verfügbarkeit von Informationen und umfangreiche Vernetzungsmöglichkeiten, birgt jedoch zugleich Risiken, wie das Verbreiten von Fehlinformation durch falsche Expert*innen (Abbildung 3) (Archila et al., 2021; Höttecke & Allchin, 2020). Sharon und Baram-Tsabari (2020) weisen darauf hin, dass eine naturwissenschaftliche Grundbildung dazu beitragen kann, Fehlinformationen im Alltag aufzudecken, aber nicht in der Art und Weise, wie sie bisher konzeptualisiert und adressiert wird. Osborne (2023) hebt in diesem Kontext ebenfalls die Bedeutung eines fundierten Verständnisses der Wissensgenese und der Eigenheiten der Naturwissenschaften hervor, welches zugleich eine Legitimation der NOS-Vermittlung im Bildungsprozess darstellt.

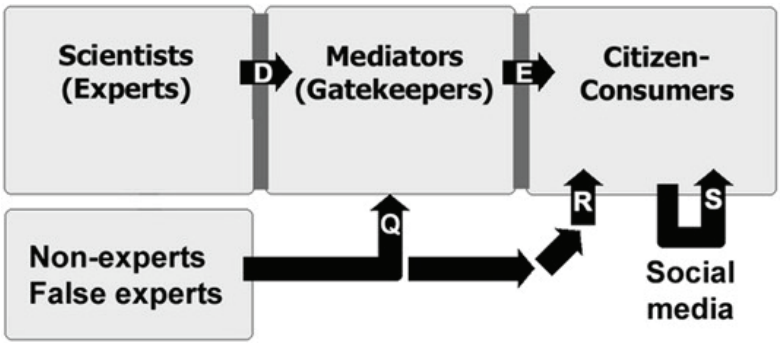


Abbildung 3: Wege der Informationsverbreitung (Höttecke & Allchin, 2020, S. 7)

Die Fähigkeit, naturwissenschaftliche Informationen von Fehlinformationen[3] zu trennen, ist heute relevanter denn je. Die JIM-Studien 2021 und 2022 zeigen auf, dass über 98 % der Jugendlichen zwischen 11 und 19 Jahren über Zugang zu einem Smartphone verfügen. 95 % beschäftigen sich dabei täglich mit Inhalten aus dem Internet, während der Anteil an analog Lesenden immer weiter abnimmt (Feierabend et al., 2021, 2022, N = 1200 in beiden Studien). Durchschnittlich verbringen Jugendliche täglich 204 Minuten online, wobei sie hauptsächlich soziale Medien nutzen und damit über drei Stunden des Tages frei publizierten Informationen ausgesetzt sind. Zunehmend sehen sie sich hier mit Fake News (56 %) konfrontiert, begegnen aber auch extremen politischen Ansichten oder Verschwörungstheorien (ebd.). Die Vodafone Stiftung Deutschland gGmbH (2020) stellte in einer repräsentativen Studie mit über 2000 Personen zwischen 14 und 24 Jahren fest, dass sich 76 % der Befragten mindestens einmal die Woche mit Fehlinformationen konfrontiert sahen, 29 % sogar täglich. Dabei spiegeln diese Zahlen lediglich die Fälle wider, in denen sich die Befragten darüber bewusst waren, dass ihnen Fehlinformationen vorlagen. In ihrer Studie zur Nachrichten- und Informationskompetenz gelangten Meßmer et al. (2021; N = 4194) zu dem Ergebnis, dass die meisten Befragten den Unterschied zwischen Desinformation, Information, Werbung und Meinung nur schwer erkennen konnten. Über 50 % erkannten eine Falschinformation nicht. 46 % der Studienteilnehmer*innen erreichten lediglich den Bereich der sehr geringen Informationskompetenz, nur 22 % erreichten hohe Werte (ebd.).

Dass Fehlinformationen nicht zu unterschätzen sind, da diese auch verhaltensrelevante Entscheidungen zur Folge haben können, zeigen agenten-basierte Modellierungen (ABM) der Universität East Anglia (Brainard et al., 2020; Brainard & Hunter,

[3] An dieser Stelle wird keine weitere Unterscheidung zwischen Fehlinformationen, welche unbewusst fälschlicherweise in Umlauf gelangen, und bewussten Falschinformationen (Fake News) vorgenommen. In jedem Fall ist die Fähigkeit zum Erkennen und Bewerten der Informationen von Relevanz.

2020). In den Computersimulationen wurden Erkenntnisse von Norovirus-Ausbrüchen extrapoliert und zeigten einen deutlichen Einfluss von Fehlinformationen auf den epidemiologischen Verlauf, da insbesondere Betroffene sich selbst und andere weniger schützen und so zur Verbreitung beitragen. Zusätzlich wurde ein Ausblick auf notwendige Maßnahmen ermöglicht, um den Einfluss von Fehlinformationen zu reduzieren. Die Datenlage legt nahe, dass die Reduktion von kursierenden Fehlinformationen von 50 % auf 40 % die negativen Effekte deutlich abmildern würde, was ebenfalls erreicht werden könnte, wenn 20 % der Bevölkerung Fehlinformation nicht teilen würden (ebd.). Da Informationen im Internet nur selten nachhaltig gelöscht werden können und jederzeit potenziell zu prüfende Inhalte hochgeladen werden, scheint nur die zweite Option, die Aufklärung der Bevölkerung, ein effektiver umsetzbarer und nachhaltigerer Schritt zu sein.

2.1.4 Pseudowissenschaften in der Gesellschaft

Der Übergang von Fehlinformationen zu gezielter Desinformation und schließlich zu Verschwörungstheorien sowie Pseudowissenschaften ist fließend. Die SARS-CoV-2-Epidemie hat verdeutlicht, wie schnell Fehlinformationen in verschwörungstheoretischen Narrativen münden können (Fuchs, 2022). Besonders Jugendliche waren während dieser Zeit mit Verschwörungstheorien, extremen politischen Ansichten und Hassbotschaften konfrontiert (Feierabend et al., 2021). Falsche Behauptungen über Zwangsimpfungen, die Angst vor implantierten Mikrochips und Aufrufe zum Protest verbreiteten sich rasant (Bundesministerium für Familie, Senioren, Frauen und Jugend, o. J.; Connect Jugendmedienzentrum et al., o. J.). Dabei geht nicht nur eine Gefahr für die eigene Sicherheit von solchen Informationen aus, auch die Gefahr einer gesellschaftlichen Spaltung nimmt zu (Krause & Gagné, 2022).

Fehlinformationen gewinnen insbesondere dann an Bedeutung, wenn wichtige Entscheidungen in der persönlichen Lebensrealität der*s Einzelnen oder der Gesellschaft getroffen werden müssen. Pseudowissenschaften erwecken hier oft den Anschein einer Wissenschaftlichkeit, die bei näherer Betrachtung nicht haltbar ist. Ein bekanntes Beispiel ist die Astrologie, die immer noch von fast 40 % der Amerikaner*innen als etwas wissenschaftlich bis sehr wissenschaftlich eingestuft wird (National Science Board, 2018). Auch in Deutschland geben fast 35 % der über 2500 Befragten der Autoritarismus-Studie 2020 an, dass das Sternzeichen sehr wahrscheinlich oder sicher einen Einfluss auf ihr Leben hätte (Decker et al., 2020; $N = 2503$). Insgesamt setzt der Esoterikmarkt schätzungsweise zwischen 15 und 20 Milliarden Euro pro Jahr in Deutschland um und die Nachfrage nach esoterischen Inhalten hat

auch durch die sozialen Medien, vor allem auf TikTok, zugenommen (Diemand, 2023). Über die sozialen Kanäle findet der Zugang meist über spirituelle Inhalte statt. Heiltees, Detox-Kuren und Ähnliches werden vermarktet und beworben. Durch diesen spirituellen Zugang ist aber auch der Schritt zur Verschwörungsmentalität nicht weit. Über 65 % der deutschen Bevölkerung geben an, latent oder stark ausgeprägt Verschwörungen in der Welt wahrzunehmen (Hellweg et al., 2020).

Heilsteinen, Tees oder Traumfängern wird oft ein esoterischer Charakter zugeschrieben, doch werden sie selten als direkte Bedrohung für Gesundheit und Gesellschaft wahrgenommen. Die daraus resultierende Bagatellisierung kann hingegen besonders im gesundheitlichen Bereich Konsequenzen haben. Beispiele für weit verbreitete alternativmedizinische Praktiken, deren Wirksamkeit bislang keine eindeutige wissenschaftliche Bestätigung gefunden haben, sind unter anderem die Chiropraktik (Ernst, 2008), die Osteopathie (Guillaud et al., 2018) oder die Homöopathie (Cucherat et al., 2000; Mathie et al., 2017). 35 % der Deutschen halten Homöopathie für wirksam, 55 % glauben zumindest teilweise an eine Wirkung (IfD Allensbach, 2023; $N = 1046$). Die Argumentation, dass Homöopathie keinen Schaden anrichten könne, wenn kein Wirkstoff vorhanden sei, ist irreführend, da die Abwendung von schulmedizinischen Therapien zugunsten homöopathischer Behandlung in schwerwiegenden Fällen zu signifikanten Gesundheitsrisiken führen kann (Ernst, 2015; Silver, 2001). Der alte Spruch *Wer heilt, hat Recht* und die Verwendung anekdotischer Evidenz tragen zur Vertiefung der Polarisation zwischen Befürworter*innen und Gegner*innen der Homöopathie bei, obwohl anekdotische Evidenzen keine wissenschaftlich fundierte Grundlage darstellen.

Alternative medizinische Praktiken erfreuen sich dennoch in Deutschland weiterhin großer Beliebtheit. Beispielsweise werden in der Geburtsvorbereitung fast ein Viertel der Geburten homöopathisch begleitet (Loytved & Schäfers, 2023). Eine Umfrage unter Hebammen ergab, dass 63 % der Befragten ($N = 309$) regelmäßig auf alternativmedizinische Methoden zurückgreifen, obwohl über 90 % der Hebammen keine formale Schulung in diesen Methoden haben (Wiebelitz et al., 2009). Auch international verwenden zwischen 65 % bis 100 % der Hebammen alternative Praktiken (Hall et al., 2012). Die Empfehlungen reichen dabei vom Einsatz von Kräutern, deren Wirksamkeit nicht ausreichend belegt sind (Simpson et al., 2001; Smeriglio et al., 2014), bis hin zum Verzehren der eigenen Plazenta (Bosco & Díaz, 2018; Coyle et al., 2015), was potenziell Risiken birgt, wie die Infektion mit Streptokokken (Buser et al., 2017; Farr et al., 2018).

Die Nachfrage und damit der Markt für alternative Medizin ist groß. Allein die Homöopathie setzte 2020 bis 2022 pro Jahr weit über eine halbe Milliarde Euro in Deutschland um, wobei nur ein kleiner Anteil der verkauften Produkte tatsächlich ärztlich verschrieben war (BAH, 2023; Bundesverband der Arzneimittel-Hersteller e.V. (BAH), 2023). Der größte Anteil entfällt auf die Selbstmedikation, was die Eigenverantwortung der Konsument*innen erhöht.

Neben der Homöopathie zählt auch die anthroposophische Medizin zu den großen alternativen Medizinrichtungen. Diese Praxis basiert auf den Lehren Rudolf Steiners und stützt sich auf Annahmen, die wissenschaftlich nicht empirisch belegbar und oft pseudowissenschaftlicher Natur sind (Scholz & Schmacke, 2022; Schöne-Seifert, 2023). Trotz fehlender Evidenz findet die anthroposophische Medizin weltweit Anklang, was sich in erheblichen Umsätzen in diesem Sektor widerspiegelt (Weleda, 2024b, 2024a). Zur Steigerung der Glaubwürdigkeit wird häufig die Nähe zur Naturheilkunde betont, obwohl die Wirkmechanismen grundsätzlich verschieden sind. Während die Naturheilkunde auf tatsächlich wirksame pflanzliche Bestandteile setzt, basieren die Wirkungsweisen der Homöopathie und Anthroposophie in weiten Teilen nicht auf der stofflichen, sondern auf einer geistig-immateriellen Ebene. Ein bekanntes Beispiel hierfür ist die Krebstherapie mit Mistelextrakten, welche nicht aufgrund ihrer Inhaltsstoffe, sondern aufgrund des Verhältnisses der Pflanze zu ihrem Wirt angewendet wird. Dieses Verhältnis wird geistig mit den Krebszellen gleichgesetzt und gemäß dem Simile-Prinzip, das besagt, Gleiches mit Gleichem zu behandeln, als alternative Krebstherapie eingesetzt (Scholz & Schmacke, 2022; Schöne-Seifert, 2023). Studien und Meta-Analysen, unter anderem vom Cochrane-Netzwerk, konnten jedoch keinen Hinweis auf eine Wirksamkeit der anthroposophischen Therapie bei Krebspatienten feststellen (Horneber et al., 2008; Seifert et al., 2022).

In Deutschland existieren zudem mehrere Lehrstühle für anthroposophische Medizin, die durch Stiftungen finanziert werden. Jene Stiftungsprofessuren, unterstützt von Organisationen wie der anthroposophischen Software-AG-Stiftung, erhöhen den Anschein einer Nähe zur konventionellen Schulmedizin und verstärken die gesellschaftliche Akzeptanz, obwohl sich Universitäten oft inhaltlich von diesen Ansätzen distanzieren (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. (AWMF) et al., 2023; Erb & Fichtner, 2023; Scholz & Schmacke, 2022; Schöne-Seifert, 2023).

Pseudowissenschaftliche oder alternativmedizinische Strömungen, welche wenig wissenschaftlich anerkannt sind, bergen nicht nur individuelle Gesundheitsrisiken,

sondern spiegeln auch eine größere gesellschaftliche Skepsis gegenüber wissenschaftlichen Erkenntnissen wider. Aktuell geben nur 56 % der Deutschen an, der Wissenschaft zu vertrauen, 44 % sind hingegen unentschieden oder vertrauen ihr nicht (Wissenschaft im Dialog gGmbH, 2023, $N = 1037$). Diese Skepsis wird durch finanzielle und institutionelle Verflechtungen oft verstärkt, was die Trennung von wissenschaftlich fundierten Methoden und solchen, die wissenschaftlich wenig anerkannt sind, erschwert.

Insgesamt zeigt sich, dass alternativmedizinische und pseudowissenschaftliche Ansätze in der Gesellschaft tief verwurzelt sind und eine klare Unterscheidung zu evidenzbasierten Methoden schwerfällt. Dies verdeutlicht die Notwendigkeit einer verstärkten wissenschaftlichen Aufklärung und Abgrenzung, um die Verbreitung von Fehlinformationen und Pseudowissenschaften zu minimieren.

2.1.5 Das Abgrenzungsproblem

Bei der Demarkation von Wissenschaft gegenüber Pseudowissenschaften stellen insbesondere die verschwommenen Grenzen ein zentrales Problem dar. Dieses Abgrenzungsproblem besteht dabei nicht nur zwischen Wissenschaft und Pseudowissenschaft, sondern generell zwischen Wissenschaft und Nicht-Wissenschaft, was Religion, Paranormales oder Verschwörungstheorien einschließt (Fasce, 2017). Zur Abgrenzung bedarf es daher zum einen einer Vorstellung, was Wissenschaft überhaupt charakterisiert, zum anderen klarer Strukturen, um sie von Nicht-Wissenschaft zu trennen.

Schon 1787 stellte Kant die Frage nach den Grenzen der wissenschaftlichen Erkenntnis. Kant argumentierte, dass menschliche Erkenntnis durch die Strukturen unseres Verstandes und die Kategorien, durch die wir unsere Welt wahrnehmen, begrenzt ist und unterschied zwischen metaphysischen Behauptungen und empirischen Erkenntnissen, die durch Erfahrung überprüfbar sind. Durch die Formulierung des Induktionsproblems von Humes (2007, nach dem Original von 1748) entstand eine Diskussion, da dieser feststellte, dass induktive Schlüsse, also Verallgemeinerungen aus einzelnen Beobachtungen, keine logische Grundlage aufweisen. Diese Annahme stellte den empiristischen Kern infrage, dass Wissen allein durch Erfahrung gesichert werden kann.

Popper (1976) lehnte die Induktion als Methode ab und schlug zur Lösung dieses Problems vor, nicht von einer Verifizierbarkeit von Aussagen, sondern von einer Falsifizierbarkeit auszugehen (Keuth, 2013). Damit etablierte er die hypothetisch-

deduktive Methode in der Wissenschaft, bei der eine Hypothese aufgestellt und diese deduktiv überprüft wird (Popper, 1935; Rosenthal, 2009). Eine Aussage wird dadurch nur dann wissenschaftlich, wenn sie prinzipiell zu widerlegen ist. Eine Aussage oder Theorie kann dieser Logik folgend nie als wahr oder richtig, sondern lediglich als bewährt gelten und unterliegt der Vorläufigkeit (Popper, 1976; Marniok, 2018).

Aus diesem Ansatz entstand eine philosophische Grundsatzdiskussion über das Wesen der Kritik in der Wissenschaft, welche hier nicht in Gänze wiedergegeben werden kann. Kuhn (1996) stimmt Popper zwar in seiner Auffassung zu, dass die Wissenschaft nicht durch die Anhäufung ewiger Wahrheiten wächst, jedoch hält er Kritik in Zeiten, die nicht als naturwissenschaftliche Revolution betrachtet werden können, für unangebracht. Er sieht den wissenschaftlichen Wandel als Paradigmenwechsel, welcher weniger rational, sondern vielmehr als mystische Konversion verläuft. Dieser Ansatz wird von Lakatos (2015) allerdings kritisiert, da er Poppers Falsifikationismus für rational begründet hält und Kuhn vorwirft, in die Irrationalität abzugleiten. Trotz dieser Diskussion hält Poppers Ansatz, nicht zuletzt wegen seines Wesens, bis heute stand, da es sich bei diesem nicht um ein Sinnkriterium, sondern um ein Abgrenzungskriterium handelt (Ströker, 1968).

Mahner (2013) hält die Falsifikation jedoch als alleiniges Abgrenzungskriterium für nicht ausreichend und ergänzt eine umfassende Liste an Kriterien zur Abgrenzung von Wissenschaft und Nicht-Wissenschaft:

1. Forschungsgemeinschaft:
   - Gibt es eine Forschungsgemeinschaft oder handelt es sich um eine lose Ansammlung von Einzelpersonen?
   - Existiert ein freier Austausch von Informationen oder gibt eine zentrale Autoritätsfigur, die Lehren an ihre Anhängerschaft weitergibt?
   - Sind eine unabhängige Forschung und Publikation möglich oder gibt es ideologische Grenzen?
2. Ontologie des Forschungsbereichs:
   - Beschäftigt sich das Feld mit konkreten Objekten oder geht es um vage Energien, Vibrationen oder übernatürliche Entitäten?
   - Geht das Feld von einer natürlichen, kausalen und gesetzmäßigen Welt aus oder werden auch übernatürliche Entitäten oder Ereignisse akzeptiert?

3. Logik und Methodologie:
   - Akzeptiert das Feld die Prinzipien der gültigen und rationalen Argumentation?
   - Werden die Prinzipien der Nicht-Zirkularität und Widerspruchsfreiheit beachtet?
   - Wird Fallibilismus (die Bereitschaft, Fehler zu akzeptieren) anerkannt oder liegt Dogmatismus vor?
   - Wird Occams Rasiermesser (die Präferenz für einfachere Erklärungen) berücksichtigt oder gibt es ungebremste Spekulationen?
4. Wichtigkeit von Testbarkeit und Kritik:
   - Wie wichtig sind Testbarkeit und Kritik im Forschungsfeld?
   - Wie wichtig ist die Unterstützung durch Evidenz?
   - Kann die Zuverlässigkeit der Methoden oder Techniken unabhängig getestet werden?
   - Haben die Theorien echte Erklärungs- oder Vorhersagekraft?
   - Sind die Theorien fruchtbar und führen zu neuen Erkenntnissen?
   - Sind die Daten reproduzierbar?
   - Gibt es Mechanismen zur Fehlereliminierung?
   - Sind die Behauptungen mit gut etabliertem Wissen kompatibel, insbesondere mit wissenschaftlichen Gesetzen?
5. Interdisziplinäre Integration:
   - Zieht das Feld Wissen und Methoden aus angrenzenden Bereichen hinzu?
   - Informiert und bereichert das Feld andere benachbarte Felder oder ist es isoliert?
6. Aktualität und Wachstum des Wissens:
   - Entstehen die Probleme des Feldes natürlich aus der Forschung oder der Theoriebildung oder werden sie künstlich geschaffen?
   - Ist der Wissensbestand des Feldes aktuell und gut bestätigt oder ist er veraltet?
   - Wächst das Wissen in diesem Bereich oder stagniert es? (vgl. ebd., S. 38 f.)

Mahner (2013) betont, dass diese Indikatoren nur eine Auswahl aller möglichen Kriterien darstellen und es unwahrscheinlich ist, dass stets alle Indikatoren erfüllt sind. Er sieht diese Kriterien vielmehr als Hinweise, um in ihrer Gesamtheit Wissenschaft von Nicht-Wissenschaft zu trennen. Sein Verständnis der Wissenschaft spiegelt die Wissenschaft als sozialen Prozess wider, der durch Offenheit, Kritikfähigkeit und

methodischer Strenge geprägt wird. Dadurch wird die Wissenschaft nicht allein inhaltlich definiert, sondern durch die Art und Weise, wie diese Ergebnisse gewonnen werden.

Ein weiterer Differenzierungsansatz ist bei Hansson (2013) zu finden, welcher wiederum zwei Wege der Abgrenzung vorschlägt. Basierend auf seiner Forschung definiert er Wissenschaft (im weiten Sinne) als die Praxis, die am zuverlässigsten und epistemisch am besten begründete Aussagen liefert, die zu einem bestimmten Zeitpunkt in den wissenschaftlichen Themenfeldern verfügbar sind (ebd.). Auf Basis dieser Definition bezeichnet er eine Aussage als pseudowissenschaftlich, wenn sie drei Kriterien erfüllt:

1. Sie bezieht sich auf ein Thema innerhalb der Domänen der Wissenschaft (Kriterium des wissenschaftlichen Bereichs).
2. Sie leidet unter einem derart schweren Mangel an Zuverlässigkeit, dass ihr nicht vertraut werden kann (Kriterium der Zuverlässigkeit).
3. Sie ist Teil einer Lehre, die den Eindruck zu wecken versucht, dass sie das zuverlässigste verfügbare Wissen auf ihrem Gebiet stellt (Kriterium der abweichenden Lehre) (vgl. ebd., S. 70 f.).

Dieser kriteriale Ansatz stellt sich im Vergleich zu Mahner (2013) deutlich reduzierter dar und bedarf einer Spezifikation des Zuverlässigkeitskriteriums. Hansson (2013) definiert daher zusätzlich einen zweiten Ansatz und führt eine Kombination verschiedener Kriterien an, welche auf Fehler zurückgehen, die üblicherweise in den Pseudowissenschaften zu finden sind:

- Glaube an Autorität,
- nicht reproduzierbare Ergebnisse,
- herausgepickte Beispiele,
- Verweigerung einer Überprüfung,
- Missachtung sich widersprechender Informationen,
- eingebaute Ausflüchte,
- Erklärungen können ersatzlos gestrichen werden (vgl. ebd., S. 72 f.).

Trifft mindestens eines dieser Kriterien auf eine Aussage zu, so ist diese nach Hansson (2013) als pseudowissenschaftlich anzusehen.

Fasce (2017) wiederum schlägt eine Verdichtung auf die Kernkriterien vor und nennt vier Kriterien, wobei mindestens eines der Kriterien 1-3 sowie Kriterium 4 erfüllt sein müssen, um eine Aussage als pseudowissenschaftlich einzustufen:

1. Bezug auf Entitäten und/oder Prozesse außerhalb des Bereichs der Naturwissenschaften,
2. Verwendung unzulänglicher Methodologie,
3. keine empirische Unterstützung,
4. Präsentation als naturwissenschaftliche Erkenntnisse (vgl. ebd., S. 468 f.).

Auch hier ist zusätzlich eine multi-kriteriale Betrachtung nötig, um diese Kriterien klarer abgrenzen zu können (Fasce, 2017).

All diese Demarkationsansätze haben gemein, dass sie lediglich als Richtlinie, denn als klare Grenze, verstanden werden müssen und ihre Verschiedenheit zeigt nachdrücklich, wie komplex sich die Frage nach einer Abgrenzung zwischen Wissenschaft und Nicht-Wissenschaft darstellt. Ein erweitertes Verständnis der Naturwissenschaften im Sinne von NOS und NOSIS ist für deren Beantwortung unumgänglich. Ein reines Fachwissen ohne epistemisches Verständnis kann dem heutigen Anspruch aus dieser Perspektive nicht gerecht werden und ist daher als notwendig, aber nicht hinreichend anzusehen. In dessen Konsequenz wird die Forderung nach einer metakognitiven Denkpraxis zur Bewertung und Abgrenzung von Informationen immer deutlicher.

Blancke und De Smedt (2013) führen die Anfälligkeit für pseudowissenschaftliche Überzeugungen vor allem auf affektive Einflüsse zurück, welchen im menschlichen Denken evolutiv begründet eine vorrangige Rolle zukommt. Die daraus resultierenden Heuristiken, mentale Abkürzungen, welche aufwendige gedankliche Prozesse vereinfachen, führen dabei zu sogenannten kognitiven Verzerrungen[4] und Fehleinschätzungen. Zusätzlich spielen weitere soziale und kulturelle Faktoren eine wichtige Rolle bei der Entscheidungsfindung und damit bei der Anwendung von Wissen.

Die Anforderungen wissenschaftlichen Denkens stehen dabei oft im Widerspruch zu den evolutionär entwickelten Denkprozessen. Wissenschaftliches Denken, welches einen hochgradig systematischen und rationalen Ansatz verfolgt, erfordert oft, dass Menschen ihre natürlichen und stark intuitiven Tendenzen überwinden. Bei-

[4] Kognitive Verzerrungen bezeichnen systematische Denkfehler, die häufig unbewusst die Urteilsfähigkeit und Entscheidungsfindung beeinflussen.

spiele hierfür stellen Bestätigungsfehler (*conformation bias*) (Mercier, 2016) oder Rahmungseffekte (*framing*) (Kühberger, 2016) dar. Wissenschaftliche Erkenntnisse zeigen sich dadurch häufig kontraintuitiv und aufgrund dessen besonders anfällig für kognitive Verzerrungen, welche in Ablehnung und Missverständnissen resultieren können.

Die Anfälligkeit für pseudowissenschaftliche Überzeugungen ist daher auch keine Folge von Ignoranz oder Unfähigkeit, sondern ist tief in den Denkstrukturen der Menschen verwurzelt. Somit ist es neben der naturwissenschaftlichen Grundbildung und dem Verständnis von NOS und NOSIS entscheidend, auch die das Denken beeinflussenden Prozesse zu verstehen. Nur so kann Wissen nicht bloß erfasst und dessen Einbettung in soziale Kontexte nachvollzogen, sondern auch im Kontext des eigenen Denkens begriffen werden.

2.1.6 Gesellschaftlich relevante Probleme im Fokus

Die Fähigkeit, naturwissenschaftliches Wissen in der Lebensrealität nutzen zu können, spielt nicht nur bei der Rezeption von Informationen eine Rolle, sondern auch beim Treffen von Entscheidungen in gesellschaftlichen Kontexten. Ganz gemäß dem *demokratischen Argument* (vgl. Kap. 2.1.2, Driver et al., 1996; Thomas & Durant, 1987) wird das naturwissenschaftliche Wissen besonders dann ausschlaggebend, wenn die gesellschaftliche Mündigkeit zur politischen Teilhabe genutzt wird. Gerade bei gesellschaftlich relevanten Problemstellungen mit naturwissenschaftlichem Bezug ist ein naturwissenschaftliches Verständnis vonnöten. Diese Problemstellungen sind in unserer Gesellschaft zunehmend allgegenwärtig (Zeidler & Kahn, 2014; Zeidler & Lewis, 2003) und besonders die Chemie weist zahlreiche Berührungspunkte mit gesellschaftlichen Themen auf, was sich in der verstärkten Forderung nach einer Förderung der Bewertungskompetenz widerspiegelt (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024b). Beispiele für solche chemiebezogenen Kontexte sind etwa die Debatte um den Klimawandel, in der insbesondere der Einfluss von Treibhausgasen diskutiert wird (Kaspar & Mächel, 2023; Rechid et al., 2023), der Atomausstieg, der eng mit Fragen der Kernkraft verknüpft ist (Schnabel, 2020), und das Problem der Plastikverwendung, das in direktem Bezug zu Mikroplastik steht (Fath, 2019). Diese Themen sind hochkomplex, entziehen sich in ihrer kontroversen Natur einer eindeutigen

Lösung und werden auch unter dem Begriff der *Socioscientific Issues*[5] (SSI) zusammengefasst (Chen & Xiao, 2021; Sadler, 2011; Zeidler et al., 2019). Um in diesen komplexen Entscheidungssituationen bestehen zu können, ist eine *funktionale naturwissenschaftliche Grundbildung*[6] vonnöten, um Lernende in ihrer Entwicklung zu informierten Mitgliedern der Gesellschaft anzuleiten (Zeidler & Kahn, 2014). Umgekehrt werden SSI als wichtige Bestandteile der Förderung einer naturwissenschaftlichen Grundbildung verstanden (Chen & Xiao, 2021; Zeidler et al., 2019), welche aber auch die Implementation von NOS einschließen sollte (Driver et al., 1996). Über den wissenschaftlichen Anspruch hinaus stellen SSI moralische und ethische Herausforderungen dar, die der Förderung einer metakognitiven Denkpraxis bedürfen, um das eigene Denken überwachen und bewerten zu können (Mestre & Cocking, 2002; Zeidler & Sadler, 2007).

SSI weisen in der Fachdidaktik vier Charakteristika auf (Zeidler, 2014; Zeidler et al., 2019). Der Unterricht mit SSI sollte demnach

1. persönlich relevante, kontroverse und unstrukturierte Probleme beinhalten, um zur Entscheidung über solche Themen wissenschaftliches, evidenzbasiertes Wissen nutzen zu können,
2. wissenschaftliche Themen mit sozialen Auswirkungen einbeziehen, um durch Dialoge, Diskussionen, Debatten und Argumentationen eine tiefere Auseinandersetzung zu fördern,
3. implizit oder explizit ethische Komponenten integrieren, welche moralisches Denken notwendig machen und
4. die Bildung des Charakters als langfristiges Ziel betonen (vgl. Zeidler, 2014, S. 699; Zeidler et al., 2019, S. 2).

Die Verwendung von SSI im Unterricht strebt dabei mehr an als nur ein besseres Verständnis von Naturwissenschaften oder das Hervorbringen von Argumentationsfähigkeiten (Sadler et al., 2007; Zeidler et al., 2019). Das übergeordnete Ziel stellt die Fähigkeit dar, komplexe Entscheidungen treffen zu können, erweist sich jedoch als

[5] Hier wird die Schreibweise Socioscientific Issues verwendet, welche von Zeidler (2014) präferiert wird, da sie eine direkte Nähe zwischen der Gesellschaft und den Naturwissenschaften impliziert. Neben dieser gibt es noch andere Schreibweisen (bspw. Socio-scientific Issues), welche auch mit einem anderen Verständnis der Terminologie einhergehen können.

[6] Der in diesem Zusammenhang verwendete Begriff der funktionalen naturwissenschaftlichen Grundbildung entspricht nicht dem Verständnis von 'Functional Scientific Literacy', wie es Bybee (2002) formulierte. Bybee (2002) versteht darunter die Anwendungsfähigkeit naturwissenschaftlicher Begrifflichkeiten und Konzepte, während Zeidler (2014) die Anwendung naturwissenschaftlichen Wissens in ethischen und moralischen Dimensionen versteht.

schwer fassbar und messbar. Um sich diesem dennoch zu nähern, entwickelten Sadler et al. (2007) das sogenannte *Socioscientific Reasoning* (SSR), welches in fünf Dimensionen die Denkpraktiken in komplexen Entscheidungssituationen greifbar machen soll:

1. *Komplexität:* Die Fähigkeit, die Komplexität der SSI zu erkennen und ihr Rechnung zu tragen.
2. *Untersuchung:* Die Informationen zu erkennen, die nicht verfügbar sind, und Wege zu finden, wie diese Informationen ermittelt werden könnten.
3. *Perspektiveinnahme:* Die Möglichkeit, ein Problem und dessen Lösungsansätze aus den Perspektiven verschiedener Positionen zu betrachten.
4. *Skepsis:* Die Fähigkeit, potenzielle Quellen von Voreingenommenheit zu identifizieren, die Informationen oder deren Präsentation beeinflussen könnten.
5. *Möglichkeiten und Grenzen der Wissenschaft:* Das Bewusstsein dafür, wie wissenschaftliches Wissen und Prozesse zur Lösung von SSI beitragen können, sowie die Erkenntnis, welche Anteile eines Problems nicht durch die Wissenschaft lösbar sind (ebd.).

Der SSR-Ansatz wurde ursprünglich mittels detaillierter Interviews durchgeführt, bei denen die Lernenden auf narrative Szenarien zu SSI reagierten. In späteren Studien wurden offene Umfragen verwendet, um eine breitere Anwendung und eine quantitative Bewertung zu ermöglichen. Diese Erhebungen forderten die Lernenden auf, ihre Gedanken schriftlich zu formulieren, wobei die Antworten anschließend nach den fünf Dimensionen des SSR bewertet wurden.

Die Studien zeigten, dass die Auseinandersetzung mit SSR in kurzen Unterrichtseinheiten (zwei Wochen zum Thema Klimawandel) keine signifikanten Verbesserungen in den SSR-Fähigkeiten der Lernenden erbrachte, eine längere Auseinandersetzung (sechs Wochen) jedoch zu messbaren Fortschritten führte (Chen & Xiao, 2021; Sadler et al., 2011; Zeidler et al., 2019).

In ihrer systematischen Übersichtsarbeit nennen Chen und Xiao (2021) verschiedene Probleme von Lernenden im Umgang mit SSI. Zum einen zeigen Lernende Schwierigkeiten mit der Unsicherheit und Uneindeutigkeit, welche die Bewertung von SSI begleiten, sowie Probleme bei der Bewertung von Quellen und Informationssuche (Ideland et al., 2011). Zum anderen zeigte sich sowohl bei Lehrenden als auch Lernenden ein Einfluss persönlicher Werte und Glaubenssätze (Cebesoy & Oztekin,

2018; Lee & Yang, 2019). Vor allem bei Lernenden spielten diese eine Rolle, wenn sich Kontexte moralisch fordernd zeigten und dazu führten, dass Entscheidungen und Diskussionen vermieden wurden. Lee und Yang (2019) plädieren daher dafür, bei der Thematisierung von SSI die moralischen und psychologischen Reaktionen der Lernenden stärker zu berücksichtigen und einen Raum zu schaffen, Werte und Überzeugungen reflektieren zu können.

Die Komplexität und moralische Uneindeutigkeiten können Gründe dafür sein, dass gerade bei SSI eine Lücke zwischen Wissen und Handeln entsteht. Diese sogenannte Knowledge-Action Gap ist insbesondere in Bezug auf den Klimawandel und Naturschutz erforscht, da hier trotz einer starken wissenschaftlichen Grundlage das Handeln der*s Einzelnen und der Gesellschaft von den Empfehlungen abweicht (Colombo et al., 2023; Dupont et al., 2024; Knight et al., 2008; Knutti, 2019; Räsänen et al., 2024; Roche et al., 2022). Dabei wird diese Lücke nicht nur durch fehlendes Wissen, sondern auch durch emotionale und affektive Barrieren verstärkt, beispielsweise wenn kurzfristiges Eigeninteresse oder eigene Gewohnheiten im Widerspruch zu besserem Wissen stehen (Colombo et al., 2023; Knutti, 2019). Zusätzlich dazu wirken sich die soziale Norm und gesellschaftliche Verflechtungen sowohl auf der Ebene der*s Einzelnen als auch auf wissenschaftlicher Ebene auf die tatsächliche Anwendung von Wissen aus (Colombo et al., 2023; Dupont et al., 2024). Entgegen der aktuellen stärkeren Gewichtung wissenszentrierter Ansätze bei der Vermittlung von Naturwissenschaften rücken damit immer mehr Argumente für eine Berücksichtigung emotional-psychologischer Faktoren in den Vordergrund, da sie bei der Bewertung von Quellen, Positionierung in SSI oder dem Treffen von Entscheidungen eine einflussreiche Rolle spielen und Ursprung kognitiver Verzerrungen sind. Angesichts dessen wird bei der Bewertung von SSI, zusätzlich zu einer naturwissenschaftlichen Grundbildung und einem Verständnis von NOS und NOSIS, die Fähigkeit immer relevanter, kritisch zu denken (Bybee, 1997; Gunn et al., 2008; Kolstø et al., 2006; Osborne, 2014; Rafolt et al., 2019; Yacoubian, 2015; Zeidler, 2014).

Komplexe Situationen, in denen naturwissenschaftliches Wissen zur Anwendung kommen soll, wie die Abgrenzung von Wissenschaft und Nicht-Wissenschaft oder die Bewertung von SSI, zeigen, dass reines inhaltliches *Verfügungswissen* oder das ergänzende Vermitteln von NOS und NOSIS nicht allein ausreichen, um Wissen in Handeln zu überführen. Obgleich ein rationales Denken angestrebt wird, ist das menschliche Denken stark intuitiv geprägt und anfällig für daraus resultierende kognitive Verzerrungen. Um diesen Herausforderungen gerecht zu werden und das menschliche Denken in seiner Gesamtheit zu adressieren, ist es nötig, die Fähigkeit

zum Kritischen Denken zu fördern, mit dem Ziel, naturwissenschaftliches Wissen, dessen Genese und Verbreitung sowie die kognitive Verarbeitung zueinander in Bezug zu setzen.

In einer immer vielschichtigeren und vernetzteren Welt erfordern komplexe kognitive Herausforderungen die Förderung komplexer kognitiver Prozesse, die über das bloße Verstehen und Anwenden von Wissen und damit auch über den Kompetenzbegriff hinausgehen. Barr und Ben-Shakhar (2018) sprechen sich daher für eine Schulung Kritischen Denkens aus, um kognitiven Verzerrungen entgegenzuwirken, die eigenen Denkprozesse zu hinterfragen, die Qualität von Argumenten einzuschätzen und Wissen zu systematisieren. Durch die Integration Kritischen Denkens in die naturwissenschaftliche Bildung könnte so die Fähigkeit gestärkt werden, wissenschaftliche Erkenntnisse nicht nur passiv aufzunehmen, sondern aktiv in den eigenen Denkprozess zu integrieren und anzuwenden.

2.2 Kritisches Denken

Kritisches Denken wird vom Rat der Europäischen Union (2017, 2018) als eine der fundamentalen Schlüsselkompetenzen für lebenslanges Lernen aufgeführt. Dabei sind Schlüsselkompetenzen definiert als

> […] diejenigen Kompetenzen, die alle Menschen für ihre persönliche Entfaltung und Entwicklung, Vermittelbarkeit, soziale Inklusion, eine nachhaltige Lebensweise, ein erfolgreiches Leben in friedlichen Gesellschaften, eine gesundheitsbewusste Lebensgestaltung und aktive Bürgerschaft benötigen (Rat der Europäischen Union, 2018, S. 7).

Kritisches Denken wird als eine zentrale Fähigkeit zur Bewertung von Informationen hervorgehoben und als notwendiger Bestandteil der *digitalen Kompetenz* betont. Ebenso spielt es eine integrale Rolle in der sogenannten *Bürgerkompetenz*, die dazu befähigt, an gesellschaftlichen Diskursen teilzunehmen und dadurch Einfluss auf eine nachhaltige Entwicklung der Gesellschaft zu nehmen. Ebenso ist Kritisches Denken ein Bestandteil der *unternehmerischen Kompetenz*, die das wertschöpfende Umsetzen von Chancen und Ideen beschreibt (ebd.).

Kritisches Denken findet sich des Weiteren sowohl explizit als auch implizit in internationalen (National Research Council, 1996) und nationalen Bildungsdokumenten, wie den Bildungsstandards des Faches Chemie (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024b, 2020),

den Standards für die Lehrerbildung[7] (Sekretariat der Kultusministerkonferenz, 2019) und den Kernlehrplänen mehrerer Länder (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2019, 2022; Sächsisches Staatsministerium für Kultus, 2022) wieder. Dort treten Teilaspekte Kritischen Denkens vorwiegend in adjektivischen Phrasen, wie *kritisch hinterfragen*, *kritisch reflektieren*, *kritisch bewerten*, *kritisch auseinandersetzten* und *kritisch umgehen* auf. Allerdings bleibt unklar, was genau unter diesen Begriffen zu verstehen ist und inwieweit sich etwa das *kritische Bewerten* vom einfachen *Bewerten* unterscheidet.

In der Naturwissenschaftsdidaktik ist Kritisches Denken als Konstrukt kaum erforscht, obwohl es dort als relevante Fähigkeit und Lernziel formuliert wird (Bailin, 2002; Byrne & Johnstone, 1987; Gunn et al., 2008; Vieira et al., 2011; Yacoubian, 2015). Zudem bleibt es oft vage und uneindeutig definiert (Rafolt et al., 2019), was eine klare Ansprache im naturwissenschaftlichen Unterricht erschwert und dazu führt, dass, wenn überhaupt, seitens der Lehrenden subjektiv geladene Definitionsansätze vermittelt werden.

Ein naheliegender Grund scheint die Komplexität zu sein, welche dem Kritisches Denken innewohnt, denn das Kritische Denken bewegt sich nicht allein auf naturwissenschaftsdidaktischer Ebene, sondern findet Beachtung in einem Kontinuum von Fachbereichen, von der Philosophie (Kant, 1787; Lipman, 1987; Paul & Elder, 2014) über die Psychologie (Dwyer et al., 2014; Greene & Yu, 2016; Halpern, 2014; D. Kuhn, 1999; Olson & Astington, 1993) und die Bildungswissenschaften (Bailin et al., 1999b; Facione, 1990; Peters, 1966; Pithers & Soden, 2000; Winch, 2003) bis zur Naturwissenschaftsdidaktik (Bailin, 2002; Gunn et al., 2008; Rafolt et al., 2019; Vieira et al., 2011; Yacoubian, 2015). Abbildung 4 zeigt eine Darstellung ebendieses Kontinuums, welches sowohl die von Kritischem Denken berührten Fachbereiche als auch die Unterscheidung der Definitionsansätze veranschaulicht.

Während die Philosophie und Psychologie Kritisches Denken als übergeordnetes Konstrukt wahrnehmen, wird in den Bildungswissenschaften und Naturwissenschaftsdidaktiken vorrangig von einer Fähigkeit gesprochen. Obschon in den folgenden Kapiteln die Schreibweise zwischen der Großschreibung *Kritischen* Denkens als Konstrukt und dem adjektivischen Gebrauch *kritisches* Denken als Fähigkeit differenziert werden könnte, wird zugunsten der Einheitlichkeit und wegen der graduellen

[7] Da es sich bei den Standards für die Lehrerbildung um den Originaltitel des Dokuments handelt, wurde dieser nicht gegendert.

Übergänge des Verständnisses die Konstruktschreibweise, sprich *Kritisches* Denken, verwendet.

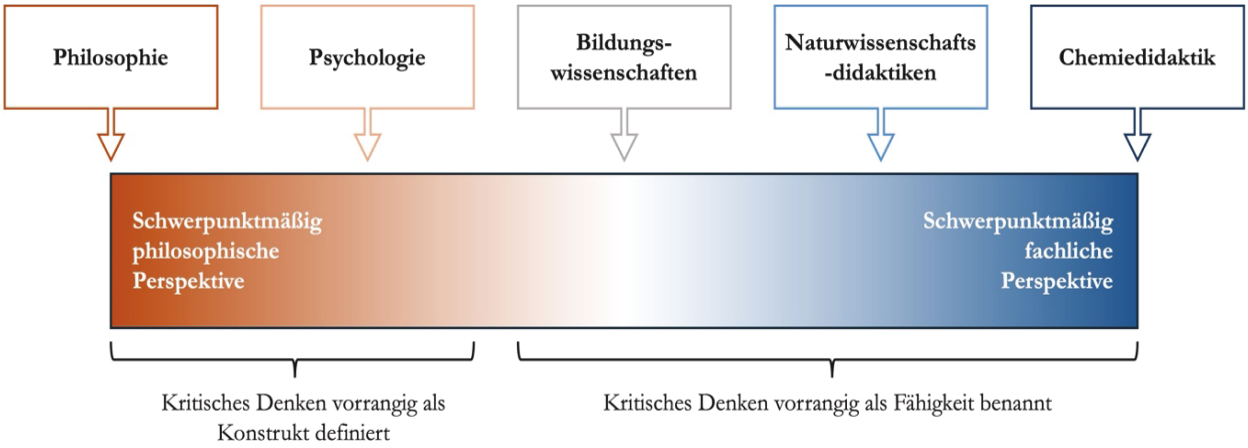


Abbildung 4: Kontinuum der Fachbereiche, welche das Konstrukt Kritischen Denkens berühren (eigene Darstellung)

Aufgrund der Ambiguität, Komplexität und Interdisziplinarität des Begriffs soll zunächst ein Überblick geschaffen werden, wie Kritisches Denken in den einzelnen Disziplinen definiert wird, bevor eine Erweiterung und Operationalisierung erfolgt. Dabei muss erwähnt werden, dass es sich um exemplarische Perspektiven handelt und kein Anspruch auf Vollständigkeit bestehen kann.

2.2.1 Naturwissenschaftsdidaktisches Verständnis und Relevanz im naturwissenschaftlichen Unterricht

Kritisches Denken wird nicht nur generell als relevante Schlüsselkompetenz angesehen, sondern auch gezielt naturwissenschaftsdidaktisch hervorgehoben. Dabei wird häufig die Forderung nach Kritischem Denken formuliert oder Anwendungsbereiche genannt, ohne tatsächlich zu definieren, was unter Kritischem Denken zu verstehen ist. Diese Forderungen können als implizite Thematisierungen Kritischen Denkens verstanden werden. Dem gegenüber stehen stärker inhaltliche Definitionsansätze, die damit als explizite Thematisierung verstanden werden können. Diese fokussieren sich jedoch schwerpunktmäßig auf die kriteriengeleitete Anwendung von Wissen und Fähigkeiten ohne Berücksichtigung emotional-psychologischer Faktoren der affektiven Ebene, welchen besonders in Entscheidungsprozessen, in denen Kritisches Denken zur Anwendung kommen soll, eine große Bedeutung zukommt (Evans, 2008).

Beispiele für eine implizite Thematisierung finden sich bei Byrne und Johnstone (1987), welche die Fähigkeit, kritisch zu denken, als implizites und explizites Ziel

naturwissenschaftlichen Unterrichts ansehen, ohne es explizit zu definieren. Sie fokussieren sich auf eine „critical-mindedness“ (ebd., S. 332), bei der sowohl kognitive Fähigkeiten als auch affektive Dispositionen eine Rolle spielen. Ihr Konzept der wissenschaftlichen kritischen Gesinnung (scietific critical-mindedness) vereint sowohl spezifische kognitive Fähigkeiten in Bezug auf wissenschaftsbezogene Kontexte als auch die Anwendungsbereitschaft. Byrne und Johnstone regen an, dieses naturwissenschaftliche Kritische Denken in komplex-verflochtenen naturwissenschaftlichen Kontexten, in denen auch soziale und ökonomische Faktoren eine Rolle spielen, zu unterrichten.

Diese Betrachtung von Fähigkeiten und Dispositionen findet sich auch in späteren Ansätzen wieder. Sinatra et al. (2014) heben hervor, dass Kritisches Denken besonders dann an Bedeutung gewinnt, wenn es darum geht, wissenschaftliche Informationen zu bewerten, vor allem, wenn wissenschaftliche Evidenz entgegen der eigenen Intuition oder Interpretation steht. Hier spielt Kritisches Denken eine zentrale Rolle bei der Auseinandersetzung mit Vorurteilen, Fehlannahmen und kognitiven Verzerrungen. In diesem Zusammenhang verweisen sie gleichzeitig auf die epistemische Kognition, das Bewusstsein, wie Wissen zustande kommt und dieses bewertet werden kann. Auch Greene und Yu (2016) erachten diese als relevant und plädieren ebenfalls für eine Vertiefung des Verständnisses von Lernenden, wie Wissen entsteht und gerechtfertigt wird.

Einen anderen Schwerpunkt setzt Bailin (2002), die ebenfalls die Förderung Kritischen Denkens als einen wesentlichen Bestandteil naturwissenschaftlicher Bildung versteht. Für sie ist Kritisches Denken hoch kontextsensibel und erfordert ein Verständnis von Konzepten, Kriterien und methodologischen Prinzipien. Kritisches Denken ist damit im naturwissenschaftlichen Unterricht auf die Einhaltung wissenschaftlicher Standards und Kriterien, wie die Genauigkeit von Daten oder die Kontrolle von Variablen, angewiesen. Während Bailin Kritisches Denken als übergeordnete Fähigkeit beschreibt, sehen Schraw et al. (2006) Kritisches Denken als eine der kognitiven Fähigkeiten selbstregulierten Lernens an und charakterisieren es als eine Reihe von Fertigkeiten, um Quellen von Informationen identifizieren und ihre Glaubwürdigkeit prüfen zu können (Abbildung 5).

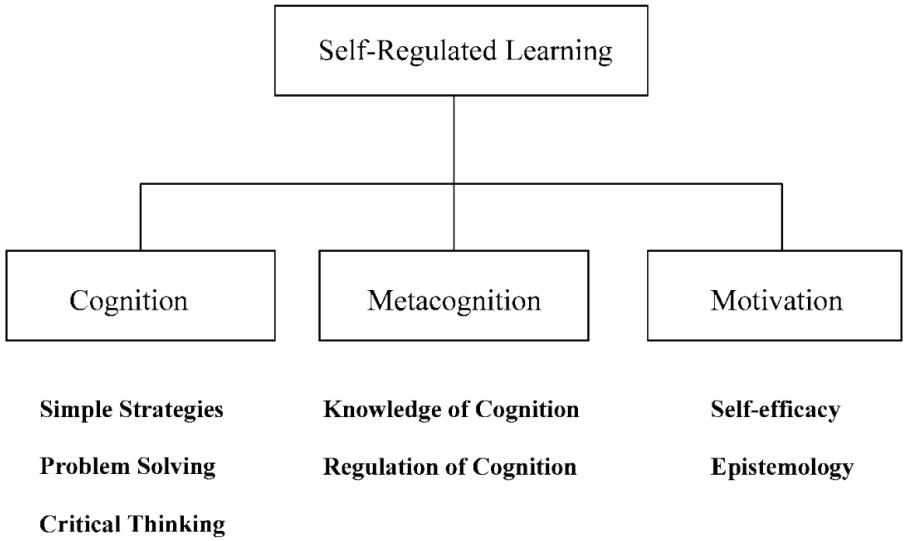


Abbildung 5: Komponenten selbstregulierenden Lernens nach Schraw et al. (2006, S. 113)

Damit steht Kritisches Denken für Schraw et al. (2006) auf einer Stufe mit dem Lösen von Problemen und dem Anwenden von Strategien und tritt nicht als übergeordnetes Konstrukt auf. Zur Förderung von Kritischem Denken und dem Lösen von Problemen diskutieren sie die Vermittlung von Strategien, was einen verstärkt inhaltsorientierten Ansatz zeigt.

Im Gegensatz dazu definieren Vieira und Kollegen Kritisches Denken angelehnt an Ennis (1989) als Kombination von Fähigkeiten und Dispositionen und heben explizit die Bedeutung Kritischen Denkens für die Lehrkräfteausbildung hervor (Franco & Vieira, 2019). Sie sehen es als notwendig an, dass Lehrkräfte Kritisches Denken unterrichten können, um Lernende zu befähigen, wissenschaftliche Methoden, Konzepte und Informationen zu hinterfragen. Sie stellen zudem einen direkten Bezug zwischen Scientific Literacy nach der Definition der National Science Education Standards (National Research Council, 1996) und Kritischem Denken her und sehen Überschneidungspunkte zwischen den Definitionen beider Begriffe (Franco & Vieira, 2019; Vieira et al., 2011; Vieira & Tenreiro-Vieira, 2016). In der Definition von Scientific Literacy der National Science Education Standards der Vereinigten Staaten (National Research Council, 1996) heißt es:

> Scientific Literacy is the knowledge and understanding of scientific concepts and process required for personal decision making, participation in civic and cultural affairs, and economic productivity. [...] It means to be able to read with understanding articles about Science in the popular press and to engage in social conversation about validity of the conclusions. [...] implies that a person can identify scientific issues underlying national and local decisions and express positions that are scientifically and technologically informed. A literate citizen

> should be able to evaluate the quality of scientific information on the basis of its source and the methods used to generate it. Scientific literacy also implies the capacity to pose and evaluate arguments based on the evidence and to apply conclusions from such arguments appropriately (ebd., S.22).

Gerade in den hier erwähnten Dispositionen sehen Vieira et al. (2011) eine Verbindung zum Kritischen Denken, insbesondere bei der Bewertung von Quellen und dem Formulieren von Schlussfolgerungen. Sie rekurrieren ebenfalls auf die Kompetenzdefinition der OECD, Wissen in Handlung zu übertragen (Organisation for Economic Co-operation and Development, 2010), und verweisen auf das Ziel der Handlungsfähigkeit in komplexen Situationen unter Mobilisierung intellektueller Ressourcen, welche für sie auch Kritisches Denken einschließt.

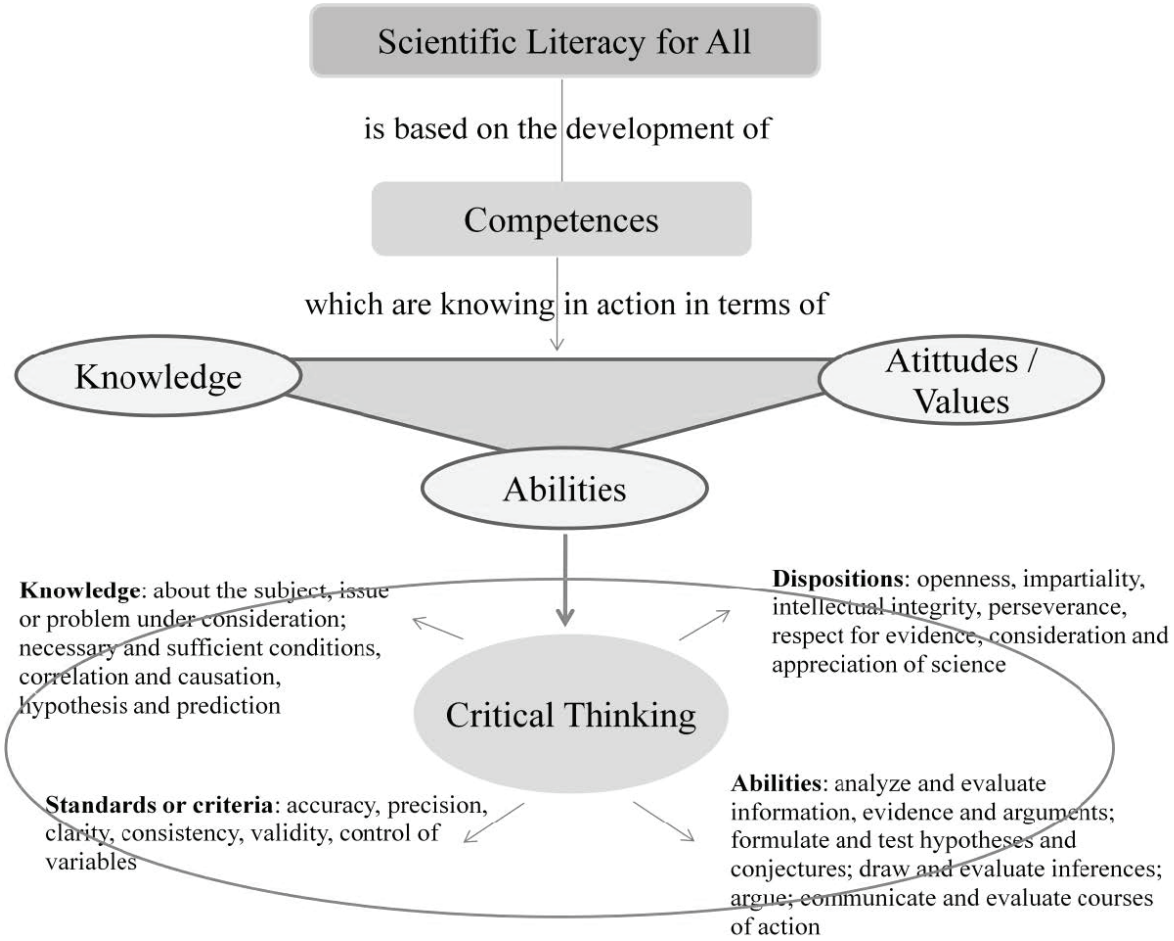


Abbildung 6: Rahmen zur Förderung des Kritischen Denkens von Lernenden im naturwissenschaftlichen Unterricht (Vieira et al., 2011, S. 50)

In ihrem Framework für die Förderung Kritischen Denkens im naturwissenschaftlichen Unterricht (Abbildung 6) konkretisieren Vieira et al. (2011) ihr Verständnis der Verbindung von Scientific Literacy, Kompetenzen und Kritischem Denken. Letzteres wird dort als eine Fähigkeit abgebildet, welche kriteriengeleitet die Anwendung

von Fähigkeiten, Wissen und den Erwerb von Dispositionen begleitet (Vieira & Tenreiro-Vieira, 2016). Zur Förderung schlagen sie die Teilnahme an Diskussionen, das Verfassen von Stellungnahmen und die Analyse wissenschaftlicher Artikel über SSI vor.

Ein ergänzender Ansatz findet sich bei Yacoubian (2015, 2018, 2020), welcher Kritisches Denken bewusst im Kontext von NOS und SSI sieht. Er schlägt ein Framework vor, welches Kritisches Denken, NOS und SSI miteinander verbindet und ein Kritisches Denken mit und über NOS anstrebt (Abbildung 7). Das Kritische Denken selbst definiert er dabei, ähnlich wie Ennis (1989), Greene und Yu (2016) oder Pithers und Soden (2000), als eine Kombination von Denkfähigkeiten und Dispositionen.

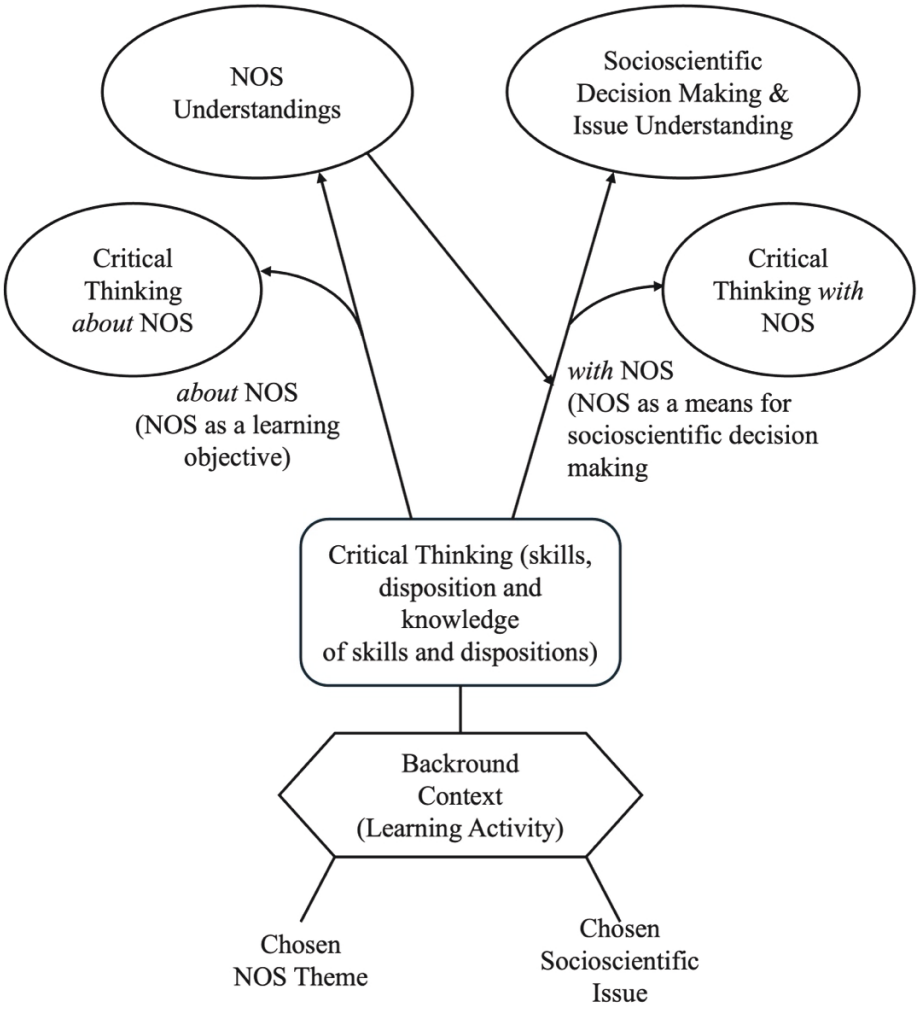


Abbildung 7: CT-NOS Framework (Yacoubian, 2015, S. 253, zur verbesserten Lesbarkeit neu gesetzt)

Dem Kritischen Denken kommt damit eine Schlüsselposition im Lernen von NOS und der Thematisierung von SSI zu, wobei diese wechselseitig bedingt sind. Das Verstehen von NOS kann in diesem Zuge im Kritischen Denken angewandt und in

der Behandlung von SSI zum Ausdruck gebracht werden. Die NOS-Charakteristika und der SSI-Kontext bilden dabei den Bezugsrahmen zur Anwendung des kritischen Denkprozesses. Gleichzeitig bildet Kritisches Denken das übergeordnete Konstrukt zur Anwendung naturwissenschaftlichen Wissens und Implementierung des NOS-Verständnisses in SSI-Kontexten.

Während Yacoubian den Fokus auf NOS und SSI legt, definieren Brambatti Guzzo und Dall'Alba (2017) Kritisches Denken im Sinne einer systematischen und rationalen Bewertung von Überzeugungen und Aussagen. Ähnlich wie bei den zuvor genannten Ansätzen umfasst auch ihre Definition sowohl kognitive Fähigkeiten zur Bewertung als auch die Disposition, diese anzuwenden. Ihre Perspektive hebt sich jedoch durch ihren starken Fokus auf kognitive Verzerrungen und die Einbeziehung der *Dual-Process* Theory von Kahneman ab, welche das Denken in ein intuitives sowie schnelles System 1 und ein langsames, aber rationales System 2 unterteilt (Kahneman, 2012). Sie argumentieren, dass ein*e kritisch Denkende*r in der Lage sein muss, kognitive Verzerrungen zu erkennen und sich von schnellen Urteilen des Systems 1 zu lösen. Eine wichtige Rolle in diesem Kontext spielt für sie die Bereitschaft, Überzeugungen zu revidieren und sich selbst zu reflektieren.

Eine stärkere inhaltliche Definition und damit explizite Thematisierung ist hingegen bei Gunn et al. (2008) zu finden. Kritisches Denken wird hier als ein intellektuell disziplinierter Prozess verstanden, welcher unter anderem die Analyse, Synthese und Bewertung von Informationen umfasst und von metakognitiver Überprüfung und Dispositionen, wie intellektueller Integrität, begleitet wird. Ferner sehen Gunn et al. (2008) im Kritischem Denken das Potenzial, zukünftige Bürger*innen zu informierten Anwender*innen von Technologie und Wissenschaft zu machen. Besondere Herausforderungen sehen sie in zukünftigen moralischen und ethischen Entscheidungen mit naturwissenschaftlichem Bezug. Gunn et al. schlagen zur Förderung dieser Denkprozesse eine explizite Instruktion der Lernenden und das Schaffen eines Bewusstseins für metakognitive Prozesse vor. Zudem sollen durch Fallbeispiele Gelegenheiten geschaffen werden, die erlernten Fähigkeiten anzuwenden.

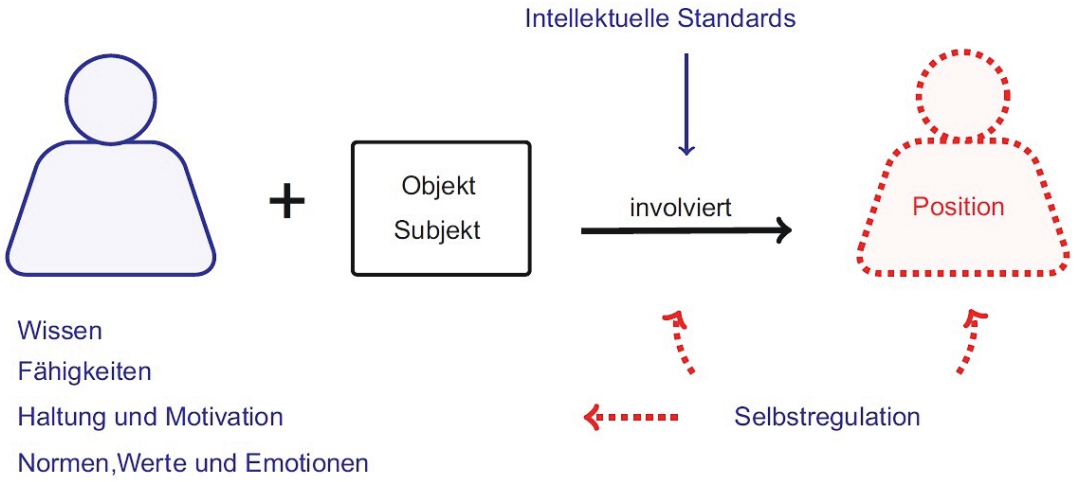


Abbildung 8: Synergiemodell des kritischen Denkens nach Rafolt et al. (2019, S. 66)

Das *Synergiemodell des kritischen Denkens* von Rafolt et al. (2019) stellt abschließend eine Annäherung und erste Synthese verschiedener Definitionsansätze Kritischen Denkens dar. Es beschreibt Kritisches Denken als kontextunabhängig und ohne normative Ausrichtung. Der Prozess selbst wird beschrieben als das Involviert-Sein der*s Denkenden mit einem Subjekt oder Objekt (eigene Haltung, Person, Problemstellung), begleitet von nichthierarchischen Charakteristika, wie Wissen und Fähigkeiten, aber auch Haltung, Motivation, Normen, Werten und Emotionen. Unter der Berücksichtigung intellektueller Standards kommt es zu einer Positionierung oder Bewertung der*s Denkenden, wobei alle Schritte stets der Selbstregulation unterliegen (ebd.) (Abbildung 8).

Rafolt et al. (2019) legen dar, dass Testungen einzelner Aspekte von Kritischem Denken nicht dessen systemischer Dimension gerecht werden und formulieren in ihrem ersten Desiderat die Frage nach dem Einfluss der Synergieeffekte auf die Qualität Kritischen Denkens. Ihr zweites Desiderat fokussiert sich auf die Erprobung der Anwendung Kritischen Denkens im Umgang mit SSI und den damit verknüpften Gelingensbedingungen. In ihrem abschließenden dritten Desiderat fordern sie die Implementation einer klaren Definition in die Lehrpläne und die Lehrkräftebildung, zur Vermittlung eines Verständnisses von Kritischem Denken und Abgrenzung von Phrasen, wie *kritisch sein* oder etwas *kritisch tun* (ebd.).

Anhand der dargestellten Definitionsansätze wird deutlich, wie unterschiedlich Kritisches Denken in den Naturwissenschaftsdidaktiken implizit und explizit thematisiert und verstanden wird: ob lediglich als Teilkompetenz des Lernens, wie bei

Schraw et al. (2006), stark kontextbezogen, wie bei Bailin (2002), mündigkeitsfördernd, wie bei Gunn et al. (2008), vernetzend, wie bei Vieira et al. (2011) und Yacoubian (2015, 2018, 2020), oder synergetisch übergreifend, wie bei Rafolt et al. (2019). Doch nicht nur innerhalb der Naturwissenschaftsdidaktiken sind Unterschiede in den Definitionsansätzen zu erkennen. Während hier Kritisches Denken schwerpunktmäßig als Kombination von Fähigkeiten und Dispositionen verstanden wird, findet es sich in anderen Fachbereichen, wie der Philosophie und Psychologie, eher als Konstrukt definiert wieder (vgl. Kap. 2.2, Abbildung 4).

Zur Erarbeitung einer umfassenden Definition zeigt sich daher die Notwendigkeit, nicht allein die naturwissenschaftsdidaktische Seite, sondern auch das historische, philosophische, psychologische und bildungswissenschaftliche Verständnis Kritischen Denkens zu beleuchten.

2.2.2 Die historische Entwicklung des Kritikbegriffs

Der Begriff der Kritik hat seinen Ursprung im griechischen Wort *krinein* (κρίνειν), das so viel wie scheiden, entscheiden oder beurteilen bedeutet. In der Antike entwickelte sich der Begriff in verschiedenen Disziplinen, wie der Medizin, dem Recht und der Philosophie. Dabei bezeichnete die *techne kritike* oder *ars critica* in der Philosophie einen Teil der Logik, der sich mit der Kunst des richtigen Urteilens befasste. In der Philologie bezeichnete die *ars critica* hingegen die Fähigkeit, Texte korrekt zu beurteilen und zu interpretieren, was heute der Hermeneutik entspricht (Röttgers, 1975).

Der moderne Kritikbegriff, wie er sich insbesondere im 18. Jahrhundert entwickelt hat, fand maßgeblich durch Immanuel Kant Eingang in die Philosophie. Kants Werke *Kritik der reinen Vernunft*, *Kritik der praktischen Vernunft* und *Kritik der Urteilskraft* sowie seine Selbstbezeichnung als *Kritizist* haben den Begriff der Kritik zu einem zentralen Bestandteil der Aufklärung gemacht. Dabei verstand Kant unter Kritik nicht nur die Beurteilung im engeren Sinn, sondern eine umfassende Untersuchung der Bedingungen und Grenzen menschlicher Erkenntnis. Die Philosophie und das Kritikverständnis Kants zielten darauf ab, die Bedingungen der Möglichkeiten von Wissen und moralischem Handeln zu erforschen und machte die Kritik zu einem systematischen Werkzeug der Aufklärung, um traditionelle Annahmen und Dogmen zu hinterfragen und zu überwinden (Kant, 1787; Röttgers, 1975; Schneiders, 1985). Bereits in diesem Grundverständnis des aufklärerischen Kritikbegriffs wird deutlich, dass nicht reines Wissen im Vordergrund stand, sondern ein ganzheitliches Verständnis der Wechselwirkungen von Wissen und Moral.

Im 19. Jahrhundert fand der Kritikbegriff in der Romantik neue Ausprägungen und wurde oft mit der ästhetischen und literarischen Beurteilung verbunden. Friedrich Schlegel verstand Kritik als schöpferische Tätigkeit, die nicht nur analysiert, sondern auch neue Bedeutungen hervorbringt. Im Gegensatz zu Kant dient die Kritik hier weniger dem strengen Urteil, sondern ist als ein Prozess der Interpretation und Erschaffung neuer Zusammenhänge zu verstehen (Röttgers, 1975). Diese Interpretation legt bereits einen systemischen Charakter nahe, da das zu Beurteilende nicht nur analysiert, sondern darüber hinaus erweitert und in neue Bedeutungszusammenhänge gebracht wird.

Im Gegensatz zu Kants und Schlegels Kritikverständnis verstand Hegel Kritik als integralen Teil des dialektischen Prozesses aus These, Antithese und Synthese. Für ihn war Kritik kein destruktiver Akt, sondern ein notwendiger Schritt zur Weiterentwicklung des Wissens und des Geistes. Dabei betonte Hegel die immanente Kritik, bei der die Widersprüche eines Systems aus sich selbst heraus überwunden werden, um zu einer höheren Synthese zu gelangen (Röttgers, 1975; Schneiders, 1985).

Im 20. Jahrhundert erlebte der Kritikbegriff eine erneute Transformation durch die Frankfurter Schule und die *Kritische Theorie*, wobei vor allem Philosophen, wie Horkheimer, Adorno und Marcuse, die Kritik zur Gesellschaftsanalyse weiterentwickelten. Diese sollte sowohl die ökonomischen, politischen und kulturellen Bedingungen hinterfragen als auch die ideologischen Strukturen, die diese stützen, um bestehende Herrschaftsverhältnisse zu überwinden. Parallel zur Kritischen Theorie entwickelte sich in der Postmoderne mit der Dekonstruktion ein weiterer Ansatz zur Kritik, welcher darauf ausgerichtet war, die Prozesse der Bedeutungskonstruktionen offenzulegen und die Machtverhältnisse, die sie stützen, zu destabilisieren (Schneiders, 1985).

In der modernen Interpretation des Kritikbegriffs kommt der *Eklektik* eine besondere Rolle zu, um unterschiedliche Denkweisen und Theorien zu kombinieren und eine umfassendere und flexiblere Analyse zu ermöglichen. Die Pluralisierung der Kritik macht deutlich, dass die Kritik heute weniger als starres Instrument des Urteils, sondern als dynamischer Prozess der Reflexion und der Integration verschiedener Perspektiven verstanden wird (ebd.).

2.2.3 Philosophische Perspektive

In der Philosophie findet vornehmlich der Prozess des Denkens selbst und seine stetige Verbesserung bei der Definition Kritischen Denkens Beachtung.

Lipman (1987) charakterisiert Kritisches Denken als selbstkorrigierendes, kriteriengeleitetes Denken, welches kontextsensibel angewandt wird. Selbstkorrigierendes Denken bezeichnet dabei nicht einfach einen metakognitiven Prozess, also das Denken über das Denken, da dies prinzipiell noch keine Selbstregulation darstellt. Er versteht darunter die aus der Beobachtung des eigenen Denkens hervorgehende Tendenz, dieses zu prüfen und zu verbessern. Dabei ist das Denken nicht ziellos, sondern kriteriengeleitet, wobei Lipman zwischen informellen und formellen Kriterien unterscheidet. Informelle Kriterien stellen dabei alltägliche Maßstäbe dar, die im täglichen Leben Anwendung finden und kontextabhängig sowie flexibel sind, wohingegen formelle Kriterien in spezifischen Kontexten akzeptiert und institutionalisiert wurden. Kritisches Denken ist dabei nicht ziellos, sondern kontextbezogen und die Denkweise an die Umstände angepasst. So sollten außergewöhnliche Umstände, Einschränkungen und der Gesamtzusammenhang berücksichtigt und auch atypische Beweise oder kontextspezifische Bedeutungen miteinbezogen werden (ebd.). Kritisches Denken wird damit als hochflexibles Denken und nicht als starres System beschrieben. Kriterien für das Denken sollten stets existieren, die Applikation dieser erfolgt aber im betrachteten Zusammenhang. Gleichzeitig findet aber nicht nur eine Bewertung des Denkgegenstandes statt, sondern stets auch eine Überprüfung und Regulierung des eigenen Denkens.

Paul und Elder (2014) strukturieren Kritisches Denken stärker und definieren es als

> […] the art of thinking about thinking while thinking to make thinking better. It involves three interwoven phases: It analyzes thinking, it evaluates thinking, and it improves thinking (ebd. S. 1).

Sie verstehen Kritisches Denken als die Kunst, sicherzustellen, dass jederzeit das bestmögliche Denken angewendet wird, und betonen die metakognitive Fähigkeit, über das eigene Denken nachzudenken. Kritisches Denken sei dabei nicht etwas, das wir unserem Denken hinzufügen, sondern eine Lebensweise, die unser gesamtes Denken durchdringt. Dabei werden drei Ebenen des Denkens unterschieden:

- Level 1: Lower Order Thinking:
  Unreflektiertes Denken, welches hauptsächlich dem intuitiven Bauchgefühl entspringt. Dieses ist geprägt von Vorurteilen und unbewussten kognitiven Verzerrungen.

- Level 2: Higher Order Thinking:
  Teilweise reflektiertes Denken, welches sich durch die Bereitschaft auszeichnet, sich mit den eigenen Denkprozessen auseinanderzusetzen. Es liegt ein Bewusstsein für die Einflüsse auf das eigene Denken vor.
- Level 3: Highest Order Thinking
  Sehr reflektiertes Denken, bei dem das Denken systematisch überwacht, analysiert und verbessert wird. Es bedarf einer willentlichen Anstrengung zur Einhaltung der intellektuellen Standards. Kognitive Verzerrungen werden bewusst wahrgenommen und Kritisches Denken routiniert in das eigene Leben integriert (ebd.).

Zudem unterscheiden Paul und Elder (2014) drei Teile des Denkens, welche berücksichtigt werden sollen: die *intellektuellen Standards*, welche auf die *Elemente des Denkens* angewandt werden sollen, um *intellektuelle Tugenden* zu entwickeln.

Die *intellektuellen Standards* umfassen:

Klarheit:	Ist das Denken deutlich und verständlich?
Genauigkeit:	Ist das Denken frei von Fehlern?
Präzision:	Ist das Denken spezifisch und genau?
Relevanz:	Bezieht sich das Denken auf das Wesentliche?
Tiefe:	Berücksichtigt das Denken alle komplexen Aspekte?
Breite:	Betrachtet das Denken verschiedene Perspektiven?
Logik:	Ist das Denken kohärent und konsistent?
Bedeutung:	Bezieht sich das Denken auf die wichtigen Aspekte des Themas?
Fairness:	Ist das Denken unparteiisch und gerecht?
Vollständigkeit:	Ist das Denken umfassend und berücksichtigt alle relevanten Aspekte?

Die *Elemente des Denkens* setzen sich zusammen aus:

Zweck:	Der Grund oder das Ziel des Denkens
Fragen:	Die Fragen, die gestellt werden, um das Denken zu lenken
Informationen:	Die Daten, Fakten und Beweise, die verwendet werden, um Schlussfolgerungen zu ziehen

Schlussfolgerungen:	Die Ergebnisse und Urteile, die aus dem Denkprozess resultieren
Konzepte:	Die Theorien, Definitionen und Modelle, die das Denken leiten
Annahmen:	Die Überzeugungen, die dem Denken zugrunde liegen
Implikationen und Konsequenzen:	Die Folgen, die sich aus den Schlussfolgerungen ergeben
Standpunkt:	Die Perspektive oder der Bezugsrahmen, aus dem gedacht wird

Die angestrebten *intellektuellen Tugenden* sind:

Intellektuelle Demut:	Das Bewusstsein für die Grenzen des eigenen Wissens und die Anerkennung der Möglichkeit, sich zu irren
Intellektueller Mut:	Die Bereitschaft, sich unpopulären oder unangenehmen Ideen zu stellen und diese zu prüfen, auch wenn sie den eigenen Überzeugungen widersprechen.
Intellektuelle Empathie:	Die Fähigkeit, sich in die Perspektiven anderer hineinzuversetzen und deren Ansichten und Argumente fair zu verstehen.
Intellektuelle Integrität:	Die Verpflichtung, den gleichen Maßstab an sich selbst anzulegen wie an andere
Intellektuelle Ausdauer:	Das Durchhaltevermögen, auch bei komplexen und schwierigen Problemen weiterhin nach der Wahrheit zu suchen
Intellektuelle Autonomie:	Die Fähigkeit, eigenständig zu denken und zu urteilen, unabhängig von äußeren Einflüssen oder Gruppendruck
Intellektuelle Unabhängigkeit:	Entwicklung eigener Gedanken und Schlussfolgerungen, statt lediglich die Meinungen anderer zu übernehmen
Intellektuelle Fairness:	Die Bereitschaft, alle relevanten Standpunkte in einer fairen und ausgewogenen Weise zu berücksichtigen, ohne voreingenommen zu sein (ebd.)

Kritisches Denken stellt für Paul und Elder (2014) ein ständiges Bestreben zur Verbesserung des eigenen Denkens unter Berücksichtigung der *intellektuellen Standards*, *der Elemente des Denkens* und *intellektuellen Tugenden* dar. Dabei handelt es sich auch hier nicht um einen einmal erlernbaren Prozess, sondern um eine Form des adaptiven

Denkens, bei dem die Formulierung der richtigen Fragen zu einem Problem maßgeblich den Denkprozess formt. Dies fassen sie prägnant zusammen mit dem Satz: „Ask questions that lead to good thinking“ (ebd., S. 137). Ebenso betonen sie die Bedeutung der Selbsterkenntnis, um kognitive Verzerrungen erfolgreich zu erkennen und zu umgehen (Paul & Elder, 2014, 2019b, 2019a).

2.2.4 Psychologische Perspektive

Auch aus psychologischer Perspektive wird Kritisches Denken oft als Metakognition beschrieben, die Fähigkeit, über das eigene Denken nachzudenken. D. Kuhn (1999) sowie Olson und Astington (1993) prägen dieses Verständnis entgegen stärker kriterialen Ansätzen, wie dem von Ennis (1987, 1989), welcher Kritisches Denken als „reasonable reflective thinking focused on deciding what to believe or do“ (ebd., S. 4) beschreibt. Ennis betont, dass Kritisches Denken sowohl Fähigkeiten als auch Dispositionen umfasst, um Aussagen bewerten zu können, und trotz eines allgemein anwendbaren Kerns domänenspezifisch ist. Im Gegensatz dazu führt D. Kuhn (1999) die Metakognition als grundlegende Denkpraxis weiter aus und unterstreicht die Notwendigkeit von Meta-Wissen, welches sie in *metakognitives Wissen*, *metastrategisches Wissen* und *epistemologisches Meta-Wissen* unterteilt. Während das metakognitive Wissen das Bewusstsein für eigene kognitive Prozesse und damit die Fähigkeit, über das eigene Denken nachzudenken, beschreibt, bezieht sich metastrategisches Wissen auf die Kenntnis verschiedener kognitiver Strategien, welche zur Problemlösung angewandt werden können. Epistemologisches Meta-Wissen umfasst abschließend das Verständnis der Natur des Wissens selbst, der Wissensgenese und seiner Kontextabhängigkeit und zeigt sich damit vergleichbar, aber weniger spezifisch, zu NOS.

Halpern (1999) sieht das metakognitive Monitoring ebenfalls als essenziellen Baustein Kritischen Denkens an, betont aber auch die Dispositionen, die dafür vorhanden sein müssen, sowie die kognitiven Fähigkeiten und Strategien, um Vorurteile zu erkennen, Argumente analysieren oder Informationen bewerten zu können. Sie versteht Kritisches Denken als umfassenden und systematischen Prozess, welcher sowohl kognitive als auch affektive Aspekte umfasst. Dieses Verständnis arbeitet sie weiter aus und bringt auch Erkenntnisse aus der Kognitionspsychologie mit ein, um kognitive Verzerrungen und Heuristiken zu erklären und für diese zu sensibilisieren (Halpern, 2014). Sie fasst ihre Arbeitsdefinition mit folgenden Worten zusammen:

> Critical thinking is the use of those cognitive skills or strategies that increase the probability of a desirable outcome. It is used to describe thinking that is purposeful, reasoned, and goal directed—the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions, when the thinker is using skills that are thoughtful and effective for the particular context and type of thinking task (ebd., S. 8).

Dabei betont sie, dass Kritisches Denken mehr ist als das Denken über das eigene Denken oder das Lösen von Problemen. Es ist ein bewusster Prozess, welcher Vorannahmen zu überwinden sucht und bestehende Werte berücksichtigt. Das *Kritische* am Kritischen Denken ist für Halpern das Bewusstsein des Denkens und das Berücksichtigen möglicher Resultate.

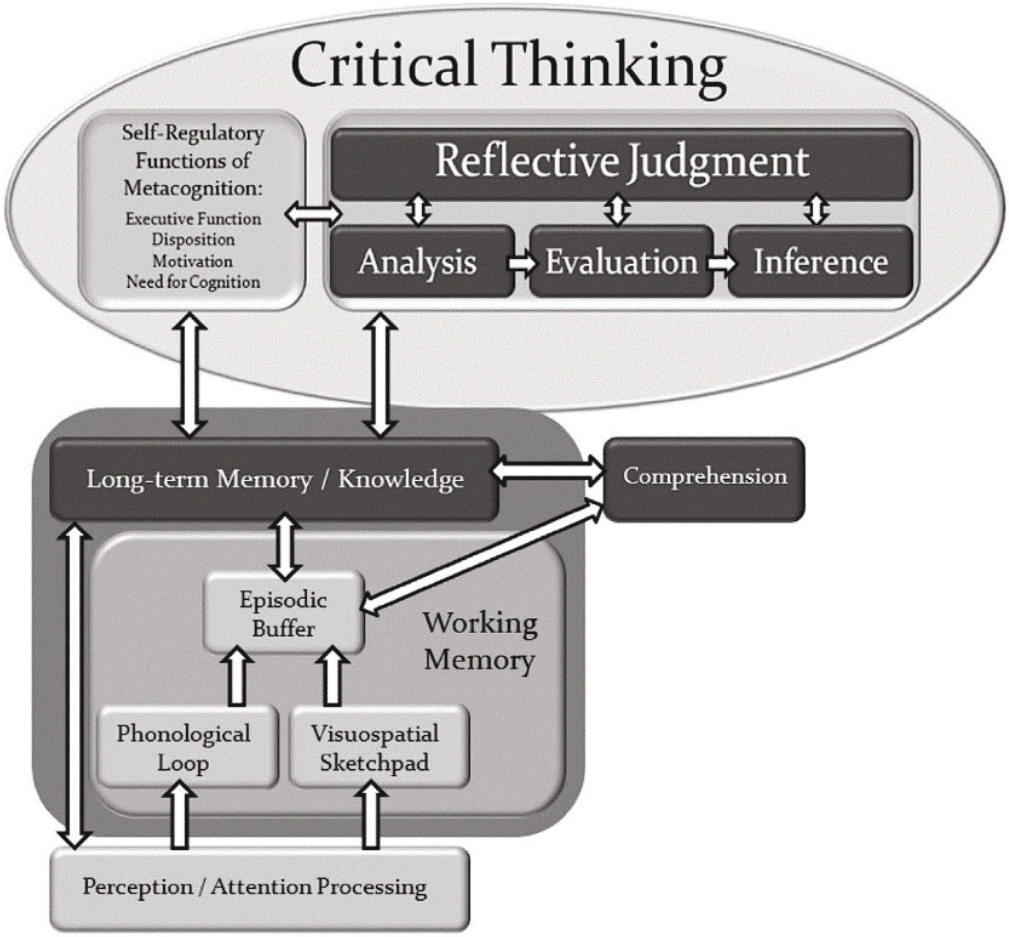


Abbildung 9: Integrative framework for critical thinking (Dwyer et al., 2014, S. 49)

Auch Dwyer et al. (2014) sehen Kritisches Denken als metakognitiven Prozess an, welcher durch gezielte, reflektierende Urteile die Wahrscheinlichkeit erhöht, eine logische Schlussfolgerung oder eine Lösung für ein Problem zu finden. Dabei stellen sie heraus, dass es nicht allein um die Anwendung kognitiver Fähigkeiten geht, son-

dern um das Überwachen der eigenen Denkprozesse währenddessen. Die Metakognition nimmt in ihrem *integrative framework of critical thinking* (Abbildung 9) daher eine relevante Vermittlerrolle ein und dient damit der Selbstregulation im Prozess des reflektierten Bewertens und während jedes Schrittes von Analyse, Evaluation und Schlussfolgerung.

Zusammenfassend lässt sich herausstellen, dass sowohl die philosophischen als auch psychologischen Definitionsansätze Kritisches Denken als ein eigenständiges Konstrukt beschreiben, das in seiner Komplexität verschiedene Elemente, Standards und Facetten umfasst. Diese Ansätze betonen die Komplexität des Konstrukts und grenzen es deutlich von einer bloßen Fähigkeit oder Einstellung ab. Aus psychologischer und philosophischer Perspektive zeigt sich Kritisches Denken als dynamischer und selbstregulierender Prozess, der kognitive und affektive Komponenten vereint und die Selbstregulation, zusätzlich zu Strategien zur Analyse und Bewertung von Informationen, als zentralen Baustein in den Vordergrund stellt.

2.2.5 Bildungswissenschaftliche Perspektive

Im Gegensatz zu den philosophischen und psychologischen Definitionsansätzen zeigt sich das Verständnis von Kritischem Denken in den Bildungswissenschaften verstärkt kriteriengeleitet und inhaltsorientiert. Bailin et al. (1999) stellen fest, dass Kritisches Denken häufig als vages Konzept verwendet wird, welches bei genauerer Definition an Übereinstimmung verliert. Sie definieren Kritisches Denken als eine normative Tätigkeit, welche angemessene Kriterien und Standards an Handlungen und Gesprochenes anlegt. In ihrem Verständnis sind dafür vorwiegend intellektuelle Ressourcen, wie Hintergrundwissen, Wissen um Schlüsselkriterien und Standards oder Heuristiken vonnöten. Sie sehen die Verantwortung der Lehrenden bei der Vermittlung von Kritischem Denken schwerpunktmäßig in der Vermittlung intellektueller Ressourcen und der Schaffung von Anwendungsgelegenheiten. Als gängiges Missverständnis über Kritisches Denken führen sie insbesondere die Tendenz an, Kritisches Denken als reine Ansammlung von Fähigkeiten oder Verfahrensschritten zu begreifen, ohne normative Standards und Kriterien zu berücksichtigen (Bailin et al., 1999a).

Pithers und Soden (2000) stellen die Bedeutung Kritischen Denkens in der Bildung heraus und argumentieren, dass dieses zwar in der modernen Bildung anerkannt ist, seitens der Lehrkräfte und Bildungseinrichtungen aber zu wenig Berücksichtigung

erfährt. Auch hier wird Kritisches Denken als ein Zusammenkommen von Dispositionen und Fähigkeiten beschrieben, darunter die Argumentationsfähigkeit, das Ziehen von Schlussfolgerungen, die Analyse von Informationen oder das logische Denken. Diese Fähigkeiten sollten dabei nicht isoliert in einem Fach, sondern übergreifend unterrichtet und durch problemorientiertes Lernen gefördert werden (ebd.).

Ein vergleichbares Verständnis von Kritischem Denken zeigen Greene und Yu (2016), welche ebenfalls die notwendigen Dispositionen und Fähigkeiten in den Vordergrund stellen. Sie verknüpfen zudem das Kritische Denken eng mit dem Konzept der *epistemischen Kognition*, einem Abwägen zwischen Dispositionen, Überzeugungen und Fähigkeiten, wie Wissen zustande kommt und wie dieses bewertet werden kann.

Winch (2003) hingegen sieht im Kritischen Denken nicht nur eine Schlüsselkompetenz für die individuelle Entscheidungsfähigkeit, sondern auch für das Funktionieren demokratischer Gesellschaften. Ihm geht es weniger um die bloße Aneignung von Wissen, sondern um die Fähigkeit, dieses reflektiert anzuwenden. Winch benennt Kritisches Denken als kritische Rationalität, welche neben der Prüfung von Argumenten und dem Schließen logischer Schlüsse auch die Angemessenheit von Handlungen in verschiedenen Kontexten bewertet. Kritisches Denken wird so nicht nur als einfacher kognitiver Prozess, sondern als soziale Praxis beschrieben, die in einem bestimmten sozialen Rahmen stattfindet und nicht nur analytische, sondern auch ethische Dimensionen beinhaltet. Damit definiert er Kritisches Denken nicht nur als Mittel zur Bewertung, sondern als Wegbereiter zu einer individuellen Autonomie durch die Befähigung, selbstständig fundierte Entscheidungen zu treffen und auch demokratisch zu partizipieren. Auch Wright (1992) sieht Kritisches Denken als relevant für die kognitive Autonomie und gesellschaftliche Teilhabe an. In der Folge spricht er sich für die Verankerung als zentrales Bildungsziel aus. Er formuliert deutlich:

> […] education, properly understood, requires critical thought by definition (ebd. S. 37).

Damit bezieht er sich auf Peters (1966), welcher als Mindestmaß für Bildung das Vorliegen guter Gründe für die eigenen Glaubenssätze nannte. In der Folge muss die Fähigkeit vorhanden sein, über die eigenen Glaubenssätze kritisch nachdenken zu können.

Eine der relevantesten und populärsten Definitionen Kritischen Denkens entstand durch die Delphi-Studie zur Definition Kritischen Denkens für Bildungszwecke von

Facione (1990). 46 Expert*innen aus Philosophie, Psychologie und Didaktik, darunter auch bereits zitierte Expert*innen wie Paul, Lipman oder Ennis, einigten sich auf eine Liste von kognitiven Fähigkeiten und affektiven Dispositionen, welche für das Kritische Denken als bedeutsam und charakteristisch eingeschätzt wurden (Tabelle 2).

Tabelle 2: Ergebnisse der Delphi-Studie von Facione (1990). Kognitive Fähigkeiten und affektive Dispositionen von Kritischem Denken (eigene Übersetzung und Darstellung)

Kognitive Fähigkeiten	**Affektive Dispositionen**
1. Interpretation: – Kategorisierung – Entschlüsselung der Bedeutung – Klärung von Bedeutungen 2. Analyse: – Untersuchen von Ideen – Identifizierung von Argumenten – Analyse von Argumenten 3. Bewertung: – Bewertung von Behauptungen – Bewertung von Argumenten 4. Schlussfolgerung: – Infragestellen von Beweisen – Annahme von Alternativen – Ziehen von Schlussfolgerungen 5. Erklärung – Ergebnisse darlegen – Verfahren rechtfertigen – Präsentation von Argumenten 6. Selbstregulation – Selbstprüfung – Selbstkorrektur	– Neugier in Bezug auf eine Vielzahl von Themen – Interesse, gut informiert zu sein und zu bleiben – Aufmerksamkeit für Gelegenheiten, Kritisches Denken anzuwenden – Vertrauen in die Prozesse der vernünftigen Untersuchung – Selbstvertrauen in die eigene Fähigkeit zu denken – Aufgeschlossenheit gegenüber unterschiedlichen Weltanschauungen – Flexibilität im Umgang mit Alternativen und Meinungen – Verständnis für die Ansichten anderer Menschen – Fairness bei der Bewertung von Argumenten – Ehrlichkeit im Umgang mit eigenen Vorurteilen, Stereotypen oder egozentrischen Tendenzen – Umsicht beim Aussetzen, Treffen oder Ändern von Urteilen – Bereitschaft, Überzeugungen zu überdenken und zu ändern, wenn eine ehrliche Reflexion darauf hindeutet, dass dies erforderlich ist

Auch hier wird deutlich, dass Fähigkeiten und Herangehensweisen den Schwerpunkt der bildungswissenschaftlichen Definition Kritischen Denkens ausmachen. Metakognitive Kompetenzen finden zwar Erwähnung, allerdings nicht so stark richtungsweisend wie in philosophischen und psychologischen Definitionsansätzen. Des Weiteren wird die affektive Ebene hier ebenfalls herausgestellt und mit ihr verbundene Dispositionen angeführt, jedoch handelt es sich schwerpunktmäßig um Eigen-

schaften, die ein*e kritisch Denkende*r erwerben sollte, und damit deskriptive Attributionen, weniger um die gezielte Berücksichtigung und Abwägung der affektiven Ebene als solche.

Den meisten naturwissenschaftsdidaktischen und bildungswissenschaftlichen Definitionen ist gemein, dass sie zwar häufig die affektive Ebene des Denkens in den Dispositionen aufgreifen, die ein*e kritisch Denkende*r benötigt, aber eher deskriptiv verbleiben. Psychologische und philosophische Ansätze stellen hingegen stärker die Metakognition in den Vordergrund und damit den Denkprozess selbst. Entgegen der Beschreibung als Werkzeug tritt Kritisches Denken dort übergeordnet und handlungsleitend auf.

Die Ziele der Naturwissenschaftsdidaktik streben nach der Förderung einer naturwissenschaftlichen Grundbildung, einem Verständnis von NOS und einer Aneignung von Strategien und Dispositionen, welche in der späteren Lebensrealität, vor allem bei komplexen Problemstellungen wie SSI, potenziell handlungsleitend sein können. Jedoch steht dem Anspruch der Rationalität diametral die stark emotional und intuitiv geprägte Natur menschlichen Denkens gegenüber (Blancke & De Smedt, 2013). Dies ist keine Schwäche, sondern entspringt einer evolutiven Grundlage, die jedoch allein durch eine Vermittlung und Förderung der rationalen Ebene nicht adressiert werden kann. In der Folge zeigt sich hier die Notwendigkeit, noch einen Schritt zurück von der Didaktik und hin zum menschlichen Denken zu machen, bevor die bisher gewonnenen Erkenntnisse wieder synthetisiert werden können.

2.3 Die Bedeutung der affektiven Ebene für das menschliche Denken

In komplexen Entscheidungssituationen, wie sie beispielsweise bei Socioscientific Issues (SSI) auftreten, zeigt sich oft, dass Wissen nicht immer zu angemessenen Handlungen führt. Dieses Phänomen, die sogenannte Knowledge-Action Gap, ist beispielsweise im Kontext des Klimawandels zu beobachten, denn obwohl eine starke wissenschaftliche Grundlage vorhanden ist, weichen das Verhalten Einzelner und der Gesellschaft häufig von den wissenschaftlichen Empfehlungen ab (Colombo et al., 2023; Dupont et al., 2024; Knight et al., 2008; Knutti, 2019; Räsänen et al., 2024; Roche et al., 2022). Ein wesentlicher Erklärungsansatz hierfür liegt in der menschlichen Entscheidungsfindung selbst, da diese nicht ausschließlich rationaler Natur ist.

2.3.1 Erste Ansätze zum Verständnis menschlichen Handelns

Einen klassischen Theorieansatz zur Erklärung menschlichen Handelns stellt die *Rational-Choice Theorie* dar. Diese nimmt an, dass Individuen auf Grundlage ihrer Wünsche und Überzeugungen rational abwägen und die Handlungsalternative bevorzugen, die den größten erwarteten Nutzen verspricht (Coleman, 1990; Downs, 1957; Elster, 2008; Voss, 2017). Elster (2008) identifiziert jedoch zwei zentrale Herausforderungen für die Rationalität in Entscheidungskontexten: *Indeterminiertheit* (Unbestimmtheit) und *Unzulänglichkeit*.

Indeterminiertheit tritt auf, wenn keine eindeutige, rationale Grundlage vorhanden ist, um eine Entscheidung zu treffen, beispielsweise wenn mehrere Optionen als gleichwertig erscheinen. In derlei Fällen tendieren Menschen dazu, zufällig oder auf Basis irrelevanter Faktoren zu entscheiden. Unzulänglichkeit zeigt sich hingegen dann, wenn die auftretenden Entscheidungen in der Praxis nicht den rationalen Annahmen entsprechen, da sich Menschen nicht immer gemäß der Theorie verhalten. Mögliche Einflussfaktoren stellen hierbei Emotionen, Impulse oder kognitive Verzerrungen dar (ebd.).

Eine Gegenperspektive zur Rational-Choice Theory bietet Elsters *Social-Norms Theory* (Elster, 1989, 2008), die davon ausgeht, dass menschliches Handeln stark durch soziale Normen geprägt wird und nicht allein ergebnisorientiert ausgerichtet ist. Soziale Normen fordern bestimmte Verhaltensweisen ein, selbst wenn diese nicht unmittelbar zu einem ergebnisorientierten oder rationalen Vorteil für die/den Einzelne*n führen. Emotionen sind an dieser Stelle besonders als Triebkräfte in der Aufrechterhaltung sozialer Normen hervorzuheben, da soziale Mechanismen, wie Lob und Tadel, in enger Verbindung zu emotionalen Reaktionen, wie Scham oder Schuldgefühl, stehen. Gleichzeitig wohnt Emotionen aber das Potenzial inne, Reflexionsprozesse anzustoßen, die eine stärkere Berücksichtigung für den Einfluss sozialer Normen ermöglichen, und damit langfristig zu besseren Entscheidungen führen (ebd.).

2.3.2 Die Dual-Process Theory

Schon Freud (2010, nach dem Original von 1900) unterschied das Bewusste und das Unbewusste und sah die Psyche als System von Kräften, welches im Gleichgewicht gehalten werden muss. Dabei maß er dem Unbewussten einen großen Einfluss auf das bewusste Denken und Verhalten bei. Diese Grundideen finden sich auch in der

von Kahneman und Tversky etablierten *Dual-Process Theory* wieder, einem der populärsten Ansätze zum Verständnis menschlichen Denkens und Handelns (Kahneman, 2012). Erste Grundzüge dieser Theorie waren bereits in ihrer *Prospect Theory* erkennbar (Kahneman & Tversky, 2013, nach dem Original von 1979), für welche Kahneman und Tversky 2002 den Wirtschaftsnobelpreis erhielten. Die Dual-Process Theory fand daraufhin weitreichende Beachtung und Anwendung zur Erklärung menschlichen Verhaltens (Chaiken, 1980; De Sousa, 2008; Epstein, 1994; Evans, 2008; Frankish, 2010; Haidt, 2008; Kahneman, 2012; Metcalfe & Mischel, 1999; Reber, 1989; Stanovich, 2008; Wegner, 1994; Wilson et al., 2000).

Diese Theorie unterteilt das menschliche Denken in zwei Systeme: das schnell agierende und intuitive *System 1* und das langsame, bewusste und rationale *System 2* (Tabelle 3).

Tabelle 3: General Features of the Two Systems (Haidt, 2008, S. 1029)

The Intuitive System	*The Reasoning System*
Fast and effortless	Slow and effortful
Process is unintentional and runs automatically	Process is intentional and controllable
Process is inaccessible; only results enter awareness	Process is consciously accessible and viewable
Does not demand attentional resources	Demands attentional resources, which are limited
Parallel distributed processing	Serial processing
Pattern matching; thought is metaphorical, holistic	Symbol manipulation; thought is truth preserving, analytical
Common to all mammals	Unique to humans over age 2 and perhaps some language-trained apes
Context dependent	Context independent
Platform dependent (depends on the brain and body that houses it)	Platform independent (the process can be transported to any rule following organism or machine)

System 1 operiert automatisch und unbewusst, ist kontinuierlich aktiv und ermöglicht schnelle Reaktionen auf Reize, ohne bewusste Aufmerksamkeitsaufwendung. Es reduziert dadurch die *Cognitive Load* (kognitive Belastung), indem viele alltägliche Entscheidungen und Abläufe automatisiert werden. System 1 basiert auf assoziativem Denken, wobei Entscheidungen auf Grundlage früherer Erfahrungen und emotionaler Eindrücken getroffen werden. Dabei werden komplexe Probleme reduziert und es kommen vereinfachte Faustregeln zum Einsatz, welche als *Heuristiken* bezeichnet werden, die eine schnelle Mustererkennung und automatische Reaktionen ermöglichen (Hoffrage et al., 2018). Diese Effizienz macht das System 1 jedoch anfällig für sogenannte *kognitive Verzerrungen*, systematische Denkfehler, die zu fehlerhaften Urteilen, Entscheidungen oder Schlussfolgerungen führen. Zudem werden

die Prozesse von System 1 durch Emotionen moduliert, welche die intuitive Entscheidungsfindung beeinflussen (De Neys, 2006a, 2006b; Haidt, 2008; Thompson et al., 2011; Frankish, 2010).

System 2 hingegen ist für langsame, bewusste und kontrollierte Denkprozesse verantwortlich. Es erfordert bewusste kognitive Anstrengung und wird aktiviert, wenn komplexe Probleme auftreten oder Entscheidungen bewusst überlegt getroffen werden müssen. Es wird im Gegensatz zu System 1 als evolutionär neuer betrachtet und als einzigartig menschlich angesehen (Frankish, 2010), da es die Fähigkeit zu abstraktem und hypothetischem Denken sowie tiefgehende Analysen, Planungs- und Problemlöseprozesse ermöglicht (Frankish, 2010; Haidt, 2008; Thompson et al., 2011).

System 1 arbeitet kontinuierlich im Hintergrund und verarbeitet eingehende Informationen, indem es fortlaufend Eindrücke, Intuitionen und emotionale Reaktionen generiert. Diese Ergebnisse werden oft ungeprüft von System 2 übernommen, sofern die Situation keine besondere Aufmerksamkeit erfordert. Erst bei komplexen Herausforderungen oder unerwarteten Ereignissen tritt System 2 bewusst in Aktion. Ein typisches Beispiel ist das Autofahren: Bei wenig Verkehr wird es weitestgehend von System 1 automatisiert gesteuert, während bei erhöhtem Verkehrsaufkommen das System 2 die Kontrolle übernimmt, was mit erhöhter kognitiver Anstrengung und schnellerer Ermüdung einhergeht. Ebenso überwacht System 2 die von System 1 generierten Vorentscheidungen, hinterfragt diese rational und passt sie gegebenenfalls an. Obwohl System 2 prinzipiell in der Lage ist, kognitive Verzerrungen zu erkennen und zu korrigieren, wird es nur dann aktiv, wenn eine Entscheidung als komplex oder wichtig wahrgenommen wird. Diese Verzögerung kann die Entstehung kognitiver Verzerrungen begünstigen und zu fehlerhaften Einschätzungen und Urteilen führen (Evans, 2018; Frankish, 2010; Haidt, 2008; Kahneman, 2012).

Ein relevanter Faktor, der auf den Übergang von System 1 zu System 2 einwirkt, ist das sogenannte *Feeling of Rightness* (FOR). Thompson et al. (2011, 2013) verstehen darunter das subjektive Gefühl, das Menschen nach einer intuitiven Entscheidung empfinden und das beschreibt, wie richtig oder zutreffend diese erscheint. Je stärker das FOR-Signal, desto weniger wird die Entscheidung bewusst überdacht und desto weniger kognitive Ressourcen werden in die Analyse der Situation investiert (Evans, 2018).

2.3.3 Intuition und kognitive Verzerrungen

Haidt (2008) hebt unter Bezugnahme auf die Dual-Process Theory die Bedeutung von Intuitionen besonders in moralischen Kontexten hervor. Sein *sozial-intuitionistischer Ansatz zur moralischen Urteilsbildung* legt nahe, dass besonders in moralischen Kontexten die Bewertungsprozesse von Intuitionen (System 1) beeinflusst werden und die Vernunftebene (System 2) nur zur nachträglichen Rechtfertigung des Urteils dient. Diese Intuitionen entspringen der affektiven Ebene und sind stark sozial, kulturell und emotional geprägt (Haidt, 2008; Sinclair, 2010).

Intuition beschreibt dabei nichts Mystisches, sondern kann vielmehr als Akkumulation von unterbewusst aufgenommenen Hinweisreizen verstanden werden, welche in passend erscheinenden ähnlichen Situationen als Grundlage der Entscheidungsbildung dienen. Der Nobelpreisträger Herbert A. Simon fasste die Intuition in seiner Studie mit Schachgroßmeister*innen mit den Worten zusammen:

> Die Situation liefert einen Hinweisreiz; dieser Hinweisreiz gibt dem Experten Zugang zu Informationen, die im Gedächtnis gespeichert sind, und diese Informationen geben ihm die Antwort. Intuition ist nicht mehr und nicht weniger als Wiedererkennen (Simon, 1992, S. 155, übersetzt in Kahneman, 2012, S. 292).

Die kognitiven Verzerrungen, die auf Basis intuitiver Entscheidungen des Systems 1 entstehen können, sind dabei vielfältig. Bekannte Beispiele hierfür sind der Bestätigungsfehler (Conformation Bias), welcher die Tendenz beschreibt, Informationen so auszuwählen und zu interpretieren, dass diese den eigenen Überzeugungen entsprechen (Mercier, 2016), oder der Anker-Effekt (Anchoring), der die Beeinflussung durch die ersten gegebenen Informationen beschreibt, welche dann als Anker die Bewertung weiterer Informationen beeinflusst (Chapman & Johnson, 2002). Darüber hinaus gibt es zahlreiche weitere Verzerrungen, wie das Überschätzen der Wahrscheinlichkeit eines Ereignisses, wenn es stark mit einem mentalen Modell oder einer Kategorie übereinstimmt (Repräsentativitätsheuristik), oder wenn entsprechende Beispiele leicht abrufbar oder emotional stark verankert sind, unabhängig von der tatsächlichen Wahrscheinlichkeit des Ereignisses (Verfügbarkeitsheuristik) (Kahneman & Tversky, 2008). Menschen tendieren zudem dazu, ihre eigenen kognitiven Verzerrungen deutlich weniger wahrzunehmen, im Gegensatz zu Verzerrungserscheinungen bei anderen, was auch als Bias Blind Spot bezeichnet wird (Pronin et al., 2002).

Die Anzahl an wissenschaftlich beschriebenen kognitiven Verzerrungen ist noch deutlich größer und liegt schätzungsweise zwischen 70 und 80 (Ghisellini & Chang, 2018). Viele kognitive Verzerrungen, Denkfehler und Abweichungen von klassischer Logik haben dabei ihren Ursprung aller Wahrscheinlichkeit nach in der Evolutionsbiologie des Menschen, seiner Entwicklung und Anpassung an Umweltbedingungen, die einen evolutionären Vorteil boten. Hinweise darauf finden sich beispielsweise in der sozialen Kognition, die das Erkennen von Absichten bei Artgenossen ermöglicht, mit dem Ziel, in komplexen sozialen Gruppen zu überleben (Cosmides & Tooby, 2008).

2.3.4 Genauere Betrachtung einzelner Einflussfaktoren

Ein großer Einflussfaktor in Entscheidung- und Bewertungsprozessen sind Gefühle, auch als *Affekt-Heuristik* beschrieben (Slovic et al., 2002). Dabei spielen Emotionen nicht nur in System-1-Prozessen eine Rolle, sondern sind aufgrund ihrer starken evolutionären Bedeutung auch in System-2-Prozessen, und damit bei rationalen Entscheidungsfindungen, von Relevanz. Diese *emotionale Inferenz* wirkt sich dadurch nicht nur darauf aus, was wir denken, sondern auch auf die Art und Weise, wie Informationen verarbeitet werden. Besonders in wissenschaftlichen Kontexten wird der Einfluss der emotionalen Inferenz häufig unterschätzt, obwohl diese eine Auswirkung auf die Entscheidungen hat, die Forschende treffen, welche Hypothesen sie verfolgen und wie sie auf Unsicherheiten oder neue Erkenntnisse reagieren. Neugier oder ein Gefühl von Dringlichkeit können sich darauf auswirken, welche Fragestellungen als relevanter eingeschätzt werden oder die Wahl von Forschungsparadigmen, -fragen oder -methoden beeinflussen (De Sousa, 2008). Tversky und Johnson (1983) stellen in diesem Zusammenhang heraus, dass Menschen sich des Einflusses der Gefühle oft nicht bewusst sind. In Konsequenz haben die mit einer Information verbundenen Emotionen häufig einen stärkeren Einfluss als der Inhalt einer Information selbst, was sich wiederum in Fehleinschätzungen und -wahrnehmungen spiegeln kann (ebd.).

Emotionen wirken sich auch bei Überzeugungsänderungen aus. Die sogenannte *kognitive Dissonanz* beschreibt das empfundene Spannungsgefühl, wenn Menschen mit widersprüchlichen Überzeugungen, Einstellungen oder Verhaltensweisen konfrontiert werden (Festinger, 1957). Das so entstehende Ungleichgewicht wird unbewusst möglichst vermieden und mittels verschiedener Bewältigungsstrategien (Coping) reduziert. Beispielsweise werden neue Informationen gesucht, die die eigene Annahme stärken, und widersprüchliche Informationen abgelehnt oder ignoriert (Harmon-

Jones & Mills, 2019), was mutmaßlich den Ursprung für den Bestätigungsfehler darstellt. Grundlage dessen ist das menschliche Bestreben, ein konsistentes Weltbild zu erhalten, weshalb eine Überzeugungs- oder Verhaltensänderung häufig mit Stress für den Menschen einhergeht und daher zunächst abgewiesen wird. Eine Änderung tritt darauffolgend oft mit der kleinstmöglichen Veränderung der eigenen Überzeugungen auf (*Minimal Change*) (Elio & Pelletier, 2008; Rott, 2008). Dieses Phänomen wird auch im Kontext der Fachdidaktik im Rahmen der epistemischen Neugier diskutiert (Berlyne, 1954; Reiners, 2022), welche als „das Bestreben des Menschen, eine Ordnung in die Erscheinungen seiner Umwelt zu bringen" (Joerger, 1976, S. 60) verstanden werden kann.

Auch irrationale Handlungen gegen besseres Wissen, philosophisch auch als *Akrasia* bezeichnet (Adler, 2008; Elster, 2008), lassen sich mit psychologischen Ansätzen erklären. Diese kommen häufig durch einen Konflikt zwischen kurzfristigen Impulsen und langfristigen Zielen zustande, wobei kurzfristigen Befriedigungen der Vorzug gegeben wird (Elster, 2008). Menschen gewichten dabei kurzfristige Belohnungen unterbewusst höher als langfristige, selbst wenn letztere höher ausfallen könnten. Neben kurzfristiger emotionaler Befriedigung können auch widersprüchliche Zielsetzungen oder Wünsche sowie unvollständige Informationen dazu führen, dass Individuen nicht die theoretisch optimale Entscheidung für sich selbst oder ihre Umwelt treffen (ebd.). Selbst der soziale Kontext und die Menschen, mit denen wir kommunizieren, haben einen Einfluss auf die Art und Weise, wie wir schlussfolgern und urteilen, was Hilton (2008) als *konversationelle Inferenz bezeichnet.*

Betrachtet man Handlungen nun als entweder *intern* oder *extern* begründet, beziehen sich interne Gründe auf die Motivationen, Überzeugungen oder Wünsche einer Person, die durch deliberative Überlegungen zu einer Handlung führen können (Williams, 2008). Diese internen Gründe sind häufig durch Emotionen beeinflusst und stark an das subjektive Motivationssystem der Person gebunden. Dabei bilden insbesondere internen Gründe verlässliche Prädiktoren für Handlungen, da sie direkt aus den individuellen Wünschen und Zielen der Person hervorgehen. Im Gegensatz dazu existieren externe Gründe unabhängig von den Wünschen, Überzeugungen oder Motivationen einer Person. Sie werden oft als objektive oder normative Gründe betrachtet, die für eine Person gelten, selbst wenn diese keine entsprechende innere Motivation besitzt, um danach zu handeln. Williams unterstreicht weiter die Bedeutung interner Faktoren, da er nur sie in der Lage sieht, eine Person tatsächlich zu motivieren (ebd.).

Auch bei der kognitiven Verarbeitung kann dahingehen unterschieden werden, inwieweit interne Faktoren auf sie einwirken. So unterscheidet Stanovich (2008) zwei Ebenen der kognitiven Kontrolle: Die *algorithmische Ebene* umfasst kognitive Prozesse, welche mit der Effizienz der Informationsverarbeitung verbunden sind, beispielsweise das Arbeitsgedächtnis und die Verarbeitungsgeschwindigkeit. Die *intentionale Ebene* hingegen bezieht sich auf Werte und Einstellungen oder Ziele, die Menschen in Bezug auf ihre Denkprozesse haben. Er argumentiert, dass gängige IQ-Tests lediglich die algorithmische Ebene erfassen, und schlägt die Begrifflichkeit der *Rationalität* vor, da die Intelligenz die anderen relevanten Dispositionen ignoriert. Rationalität erfordert hingegen sowohl die Fähigkeit, Informationen zu verarbeiten (*algorithmisches Niveau*), als auch die Fähigkeit, epistemische Ziele zu setzen und die Bereitschaft zu haben, Überzeugungen auf der Basis neuer Beweise anzupassen (*intentionales Niveau*). Eine weitere Unterscheidung nimmt er zwischen der *epistemischen Rationalität* und *praktischen Rationalität* vor. Die epistemische Rationalität bezieht sich auf die Fähigkeit, gute Überzeugungen zu formen, Beweise zu hinterfragen und Argumente zu berücksichtigen, wohingegen die *praktische Rationalität* sich auf die Fähigkeit bezieht, effizient auf Grundlage der eigenen Ziele und Überzeugungen zu handeln, um seine Ziele zu erreichen. Beide Ebenen, die algorithmische und die intentionale, also sowohl die Möglichkeit der Verarbeitung als auch die Bereitschaft, determinieren in Summe die Rationalität des Menschen, weswegen auch intelligente Menschen irrational handeln können (ebd.).

Beide Ansätze, die Unterscheidung von Handlungen nach Williams (2008) sowie die Unterscheidung von kognitiven Verarbeitungsprozessen nach Stanovich (2008), unterstreichen damit nochmals den Einfluss interner affektiver Aspekte, wie Emotionen, Motivationen, Überzeugungen, Wünsche oder Ziele.

2.3.5 Theory of Planned Behaviour

Eine weitere Theorie, die an dieser Stelle berücksichtigt werden soll, ist die Theorie des geplanten Verhaltens (Theory of Planned Behavior, TPB) von Ajzen (1991), die bereits in der Biologiedidaktik Anwendung findet. Dort wird sie beispielsweise zur Untersuchung beeinflussender Faktoren auf die Intentionen von Lehramtsstudierenden eingesetzt (Aptyka & Großschedl, 2022; Weber et al., 2020; Weber & Fiebelkorn, 2019). Die Theorie des geplanten Verhaltens stellt eine Weiterentwicklung der Theorie des überlegten Handelns (Theory of Reasoned Action, TRA) (Ajzen & Fishbein, 1975) dar, welche versucht, menschliches Handeln zu erklären und vorauszusagen. In der ursprünglichen TRA wird davon ausgegangen, dass Menschen ihr Verhalten

basierend auf ihren Intentionen und Einstellungen steuern, wenn sie die volle Kontrolle über die Durchführung dieses Verhaltens haben. Die TPB erweitert die TRA dahingehend, dass sie den zusätzlichen Faktor der wahrgenommenen Verhaltenskontrolle einführt. Diese bezieht sich auf die Einschätzung einer Person, ob und in welchem Maße sie sich selbst in der Lage sieht, das Verhalten unter den gegebenen Umständen tatsächlich umzusetzen. Diese Wahrnehmung determiniert, inwiefern die beabsichtigte Handlung letztlich erfolgt (Ajzen, 1991).

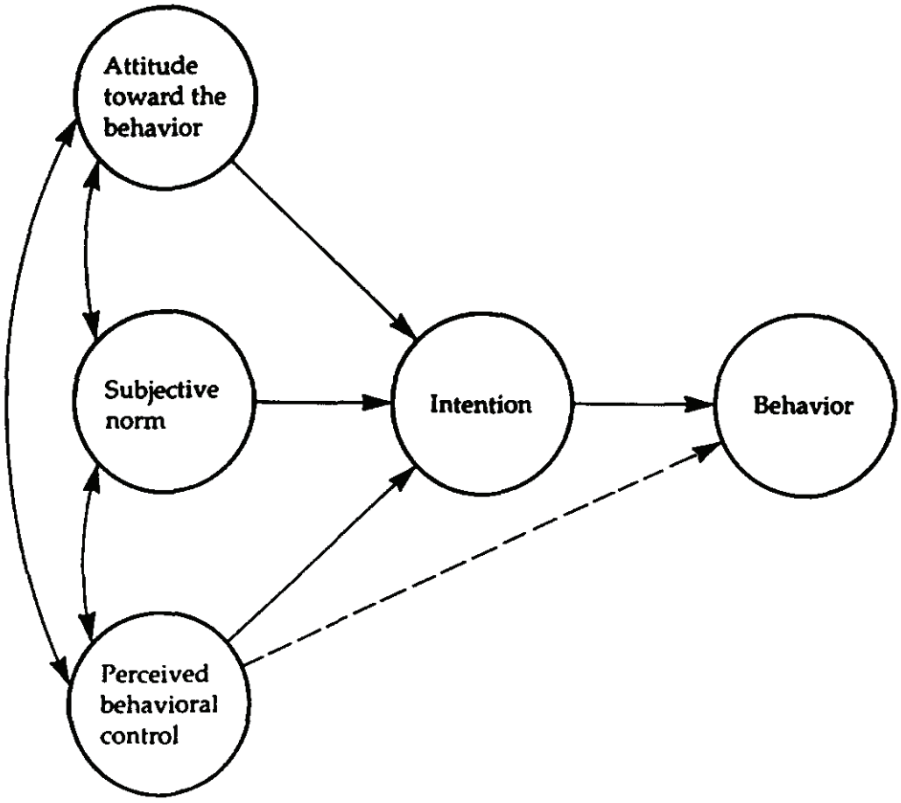


Abbildung 10: Theorie des geplanten Verhaltens (Ajzen, 1991, S. 182)

Die Theorie des geplanten Verhaltens beschreibt drei wesentliche Einflussfaktoren auf die Verhaltensintention, also die Motivation, ein Verhalten zu zeigen (Abbildung 10). Die Einstellungen gegenüber einem Verhalten beziehen sich auf damit verbundene positive oder negative Bewertungen, während die subjektiven Normen als soziale Einflüsse sowie den daraus resultierenden wahrgenommenen Druck, ein Verhalten auszuführen oder zu unterlassen, verstanden werden können. Die wahrgenommene Verhaltenskontrolle beschreibt, wie schwer oder leicht eine Person das Ausführen eines bestimmten Verhaltens einschätzt, und berücksichtigt sowohl die eigenen Fähigkeiten als auch externe Bedingungen. Diese wahrgenommene Verhaltenskontrolle beeinflusst dabei nicht nur die Intention, ein Verhalten zu zeigen, sondern auch direkt die Wahrscheinlichkeit, dass ein Verhalten ausgeführt wird. All diese

Faktoren, die Einstellung, subjektive Normen und die wahrgenommene Verhaltenskontrolle, wirken dabei nicht isoliert, sondern formen gemeinsam die Verhaltensintention, welche den stärksten Prädiktor für tatsächliches Verhalten darstellt (ebd.).

2.3.6 Der Einfluss der affektiven Ebene auf die Anwendung von Wissen

Mit Blick auf die verschiedenen Theorien der Kognitionspsychologie zeigt sich, wie viele Faktoren das menschliche Denken beeinflussen, ohne dass dies bewusst wahrgenommen wird. Die affektive Ebene, das intuitive System 1, wird dabei vielschichtig von verschiedensten internen und externen Hinweisreizen gespeist, wie den Emotionen, Motivationen, Einstellungen, Werten, Normen, Wünschen, Zielen oder der sozialen Norm. Wissen kann zwar erworben und Fähigkeiten erlernt werden, deren Umsetzung ist aber zusätzlich von vielen anderen unbewussten Faktoren mitabhängig.

Dies erklärt auch die oft beobachtete Diskrepanz zwischen Wissen und Handeln, wie sie im Kontext des Klimawandels auftritt. Auch wenn das rationale System 2 eine fundierte Entscheidungsgrundlage bietet, sind es oft die unbewussten, emotionalen Prozesse, welche die Handlung tatsächlich lenken. Kognitive Verzerrungen, die durch intuitive Urteile und Heuristiken hervorgerufen werden, können zu falschen Schlussfolgerungen führen, selbst wenn das verfügbare Wissen auf soliden wissenschaftlichen Erkenntnissen basiert. Menschliches Denken und Handeln muss dabei immer als ein Zusammenspiel aus rationalen, emotionalen und sozialen Einflüssen betrachtet werden. Nur wenn diese vielschichtigen Faktoren im Blick behalten werden, lassen sich komplexe Entscheidungssituationen und das menschliche Verhalten in seiner Gesamtheit verstehen.

Zusammenfassend wird deutlich, dass menschliches Denken nicht isoliert auf der rationalen Ebene stattfindet. Häufig geben unbewusste, affektive Prozesse den Ausschlag für Handlungen, wenngleich das verfügbare Wissen eine andere Entscheidung nahelegt. Diese affektiven Einflüsse verdeutlichen, dass ein rein rationaler Ansatz zur Wissensvermittlung nicht ausreicht, um angemessenes Handeln in komplexen Entscheidungssituationen, wie bei Socioscientific Issues, zu fördern. Es erscheint daher notwendig, die affektiven und sozialen Einflüsse in Bildungsprozesse einzubeziehen, um Kritisches Denken zu stärken und die Diskrepanz zwischen Wissen und Handeln zu verringern.

3 Konzeptioneller Teil - Synthese

Mit Blick auf die vorangegangenen Kapitel erfolgt in diesem Abschnitt eine Synthese der bisherigen Erkenntnisse. Zunächst werden die Forschungslücken identifiziert und die zugrundeliegenden Leit- und Forschungsfragen formuliert. Anschließend wird die Bedeutung der affektiven Ebene in den Naturwissenschaften herausgestellt und, basierend auf den identifizierten Schwerpunktclustern bestehender Definitionen, eine eigene Neudefinition des Kritischen Denkens entwickelt. Darauf aufbauend erfolgt eine Abgrenzung dieser Neudefinition von bestehenden Konzepten. Abschließend wird die Relevanz des Kritischen Denkens in den Naturwissenschaftsdidaktiken hervorgehoben sowie dessen Beitrag zur Bildung verdeutlicht.

3.1 Ableitung der Forschungsfragen

Die Erkenntnisse der vorangegangenen Kapitel (s. Kap. 2) bekräftigen die zentrale Bedeutung Kritischen Denkens nicht nur allgemein, sondern auch in der Naturwissenschaftsdidaktik, insbesondere der Chemiedidaktik. Trotz seiner klaren Benennung als Schlüsselkompetenz und der Betonung seines Einflusses auf andere zentrale Kompetenzbereiche, wie der digitalen Kompetenz (Rat der Europäischen Union, 2017, 2018), der expliziten Forderung nach der Implementierung Kritischen Denkens im naturwissenschaftlichen Unterricht (Bailin, 2002; Byrne & Johnstone, 1987; Gunn et al., 2008; Rafolt et al., 2019; Vieira et al., 2011; Yacoubian, 2015, 2020) sowie seiner expliziten Nennung als Bildungsziel in internationalen Bildungsstandards (National Research Council, 1996) und impliziten Aufgriffen in nationalen Bildungsrichtlinien für Lernende und Lehrende (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2022; Sächsisches Staatsministerium für Kultus, 2022; Sekretariat der Kultusministerkonferenz, 2019; Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024b), etwa in der Forderung des kritischen Reflektierens, verbleibt das Konzept des Kritischen Denkens weiterhin ambig definiert und schwer operationalisierbar.

Daraus erwächst ein Spannungsfeld, welches in dieser Arbeit als Forschungslücke identifiziert wird.

Ziel dieser Arbeit ist es, langfristig zur Förderung des Verständnisses und der Anwendung Kritischen Denkens in der Naturwissenschaftsdidaktik, insbesondere der Chemiedidaktik, beizutragen.

Daraus ergibt sich die folgende das Projekt umspannende Leitfrage:

Leitfrage:

Wie kann Kritisches Denken im naturwissenschaftlichen Unterricht nachhaltig gefördert werden?

Zur Annäherung an diese Leitfrage besteht das Forschungsvorhaben aus einem konzeptionellen Teil, welcher Kritisches Denken zu definieren sucht, und einem empirischen Teil, welcher sich der Operationalisierung des Konstruktes annimmt. Hieraus ergeben sich zwei separate Forschungsfragen, wobei die Beantwortung der konzeptionellen Forschungsfrage die Grundlage für die Behandlung der empirischen Forschungsfrage darstellt:

Konzeptionelle Forschungsfrage:

Wie kann Kritisches Denken als Konstrukt definiert werden?

Empirische Forschungsfrage:

Wie kann Kritisches Denken als Konstrukt im Kontext der Chemiedidaktik operationalisiert werden?

Zielsetzung der Arbeit ist damit die Definition des Konstruktes und die Entwicklung eines Testinstrumentes, welches geeignet ist, Kritisches Denken greifbar und zugänglich zu machen, um langfristig zu einer Förderung Kritischen Denkens in der Chemiedidaktik beizutragen.

Um eine fundierte Operationalisierung zu ermöglichen, ist jedoch zunächst eine klare und explizite Definition des Kritischen Denkens sowie die Legitimierung der Berücksichtigung der affektiven Dimension in den Naturwissenschaften notwendig.

3.2 Die Bedeutung der affektiven Ebene in den Naturwissenschaften

Die Naturwissenschaften unterliegen dem Anspruch von Neutralität, Objektivität, Transparenz und Nachvollziehbarkeit. Affektive Einflüsse sollten in den Naturwissenschaften weitestgehend reduziert sein, um möglichst objektive Ergebnisse in der Forschung zu gewährleisten. In der Konsequenz legen die Naturwissenschaftsdidaktiken den Schwerpunkt in der Vermittlung primär auf kognitive Aspekte und die

Förderung von Wissen und Fähigkeiten. Dennoch können sich die Naturwissenschaften vom Einfluss der affektiven Ebene nicht freisprechen, denn obgleich des hohen rationalen Anspruchs der Naturwissenschaften ist es der Mensch, mit all seinen Eigenschaften und Einflüssen, der sie gestaltet, ausführt und weiterentwickelt. Aufgrund dessen ist die Wissenschaft, so wie der Mensch selbst, von Subjektivität, Verzerrungen und Interpretationen beeinflusst.

Menschen nehmen die Welt durch den Blickwinkel der Muster und Regeln wahr, insbesondere in Kontexten, in denen Entscheidungen oder Bewertungen auf Grundlage unvollständiger und unsicherer Informationen getroffen werden (vgl. Kap. 2.3). Dies ist beispielsweise bei induktiven Schlüssen der Fall, da jene keine logische Gewissheit bieten, was sie anfällig für kognitive Verzerrungen, wie Übergeneralisation, dem Sehen von Mustern, wo keine sind, oder Voreiligkeit, macht. Bezogen auf den empirischen Kern der Naturwissenschaften resultiert dieser Umstand im sogenannten Induktionsproblem, welchem Popper mit der Falsifizierbarkeit von Hypothesen und der hypothetisch-deduktiven Methode begegnet (Popper, 1935; Rosenthal, 2009) (vgl. Kap. 2.1.5).

Popper hebt in diesem Zusammenhang hervor, dass wissenschaftliche Theorien, welche für den Charakter der Wissenschaft gemeinsam mit der Empirie unabdingbar sind, nie durch eine Vielzahl von Beobachtungen verifiziert, sondern nur falsifiziert werden können (Reiners, 2022; Ströker, 1968) (vgl. Kap.2.1.5). Im hypothetisch-deduktiven Verfahren werden Hypothesen aufgestellt, aus denen deduktiv überprüfbare Vorhersagen abgeleitet und anschließend empirisch getestet werden. Solange diese Vorhersagen mit den Beobachtungen übereinstimmen, gilt die Hypothese als vorläufig bestätigt. Treffen jedoch Beobachtungen ein, die im Widerspruch zur Hypothese stehen, muss diese verworfen oder angepasst werden (Popper, 1935; Rosenthal, 2009).

Obwohl das hypothetisch-deduktive Verfahren darauf abzielt, potenzielle Verzerrungen zu minimieren, ist die Wissenschaft dennoch nicht vollkommen davor gefeit. Dies liegt unter anderem daran, dass auch im hypothetisch-deduktiven Vorgehen induktive Elemente nicht vollständig vermieden werden können (Rosenthal, 2009). Kognitive Verzerrungen können bereits bei der Formulierung der Hypothesen auftreten, etwa durch den Bestätigungsfehler (Confirmation Bias), oder die Interpretation der empirischen Ergebnisse kann durch Voreingenommenheit sowie die Tendenz zur selektiven Wahrnehmung beeinflusst werden. Die Abhängigkeit vom Peer-Review-Prozess und die Bedeutung von Transparenz und Reproduzierbarkeit sollen

diese Verzerrungen zwar reduzieren, können aber nicht alle Einflüsse vollständig eliminieren. So bleibt die wissenschaftliche Praxis, trotz der systematischen Methodologie, anfällig für kognitive Verzerrungen, die das wissenschaftliche Urteilsvermögen beeinflussen können.

Neben den internen Herausforderungen der Wissenschaft prägen auch externe Faktoren die wissenschaftliche Praxis. So tragen der hohe Publikationsdruck, die Tendenz, überwiegend positive Ergebnisse zu veröffentlichen (Publikationsbias), Finanzierungsstrukturen und Interessenkonflikte maßgeblich dazu bei, dass wissenschaftliche Arbeit auf individueller und kollektiver Ebene beeinflusst wird, was in einigen Fachbereichen zu Problemen bei der Reproduzierbarkeit von Ergebnissen führt (Bausell, 2021).

Ferner spielen politische Rahmenbedingungen eine entscheidende Rolle für die Förderung oder Behinderung bestimmter Forschungsbereiche. Ein Beispiel hierfür ist die Reduktion der Finanzmittel für die Klimaforschung in den USA während der Amtszeit von Präsident Trump (The Independent, 2018). Zusätzlich verdeutlicht die Rolle des Zufalls im naturwissenschaftlichen Erkenntnisprozess, wie bei der Entdeckung des Teflons, der Röntgenstrahlung, des Newtonschen Gravitationsgesetzes, der Entdeckung der DNA oder des Penicilins, dass wissenschaftliche Erkenntnisse nicht immer gerichtet entstehen müssen, sondern von zusätzlichen Faktoren abhängig sind (Roberts, 1989), was dem Mythos der universellen Methode der Naturwissenschaften entgegensteht (McComas, 2020).

Die Thematisierung von *Nature of Science* macht bereits einen Schritt, um die Rolle des Menschen im wissenschaftlichen Prozess zu berücksichtigen (Driver et al., 1996; Lederman, 1992; McComas & Olson, 1998; Osborne et al., 2003). Aspekte wie die soziokulturelle Eingebundenheit und Kreativität verdeutlichen einzelne Facetten der Wechselwirkungen von Menschen und Wissenschaft, insbesondere auf gesellschaftlicher Ebene (vgl. Kap. 2.1.2).

McComas (2020; McComas et al., 2020) benennt die Bedeutung des Menschen und seine Subjektivität im Erkenntnisgewinnungsprozess, rekurrierend auf NOS, als das *Human Element of Science* (Abbildung 1). Diesen Ansatz aufgreifend lässt sich aber argumentieren, dass das Human Element of Science nicht nur bei der Genese naturwissenschaftlicher Erkenntnisse, sondern auch bei deren Rezeption vorhanden ist. Erkennbar wird der menschliche Einfluss dort beispielsweise in Verzerrungen, welche auf die kognitive Dissonanz zurückgehen, oder in einer Anfälligkeit für pseudowissenschaftliche Überzeugungen trotz gegenteiliger Evidenz. Der Grund hierfür

liegt wiederum in der affektiven Ebene, da die Anforderungen wissenschaftlichen Denkens oft in Widerspruch zu den evolutionär entwickelten Denkprozessen stehen und stark von diesen beeinflusst sind (Blancke & De Smedt 2013).

Die Relevanz der affektiven Dimension bei der Rezeption naturwissenschaftlicher Erkenntnisse wird besonders deutlich im Rahmen des NOSIS-Konzepts, das NOS im Kontext des digitalen Zeitalters betrachtet (vgl. Kap. 2.1.3, Höttecke & Allchin, 2020) (Abbildung 2). Hier werden neben den rationalen Faktoren auch explizit affektive Aspekte in Form der kognitiven und sozialen Filter bei der Rezeption naturwissenschaftlicher Erkenntnisse thematisiert sowie auf kognitive Verzerrungen und Heuristiken rekurriert. Neben diesem expliziten Verweis muss festgehalten werden, dass nicht nur dieser Teil, sondern der gesamte wissenschaftliche Prozess nicht frei von affektiven Faktoren ist, denn bereits bei der Beobachtung und der Interpretation von Informationen unterliegt der Mensch emotionalen und sozialen Einflüssen. Der Bestätigungsfehler (Conformation Bias) stellt in diesem Zusammenhang eine der häufigsten Verzerrungen in wissenschaftlichen Arbeiten dar. Er leitet unbewusst dazu an, nur die Daten im Forschungsprozess zu berücksichtigen, welche die bestehende Hypothese stützen (Mercier, 2016), was zu verfälschten Ergebnissen und falschen Schlussfolgerungen führen kann. Darüber hinaus kann die Verfügbarkeitsheuristik dazu beitragen, dass mental präsente Studiendesigns oder Interpretationen unbewusst Einfluss auf die Planung des eigenen Forschungsvorhabens nehmen (Kahneman & Tversky, 2008). Zudem könnte die soziale Norm Erklärung für den sogenannten Publikationsbias bieten, indem die Erwartungshaltung der wissenschaftlichen Gemeinschaft sowie die Angst vor Ablehnung dazu führen, dass vornehmlich positive Ergebnisse veröffentlicht werden (Bausell, 2021). Nicht zuletzt spielen Emotionen eine bedeutende Rolle im wissenschaftlichen Prozess und beeinflussen diesen auf verschiedenen Ebenen.

Emotionen sind dabei nicht per se negativ oder positiv zu bewerten, da sie unveränderlich Teil der menschlichen Natur sind und auch zu Kreativität und Innovation führen können, was einen wesentlichen Teil zum Fortschritt der Wissenschaft beiträgt. Dennoch birgt der Mangel an Bewusstsein für diese emotionale Inferenz Risiken, denn diese wirkt sich durch Gefühle, wie Neugier, Stress, Angst oder Dringlichkeit, unbewusst auf viele wissenschaftlich relevante Schritte aus, beispielsweise auf die Wahl der Forschungsfrage- und Paradigmen, den Umgang mit Unsicherheiten oder Reaktionen auf neue Erkenntnisse (vgl. Kap. 2.3, De Sousa, 2008).

Diese affektiven Einflüsse und kognitiven Verzerrungen sind nicht nur theoretische Überlegungen, sondern spiegeln sich in vielen historischen und aktuellen wissenschaftlichen Prozessen wider. Ein prägnantes Beispiel hierfür ist der Fall von Daniel Shechtman, dessen Entdeckung der Quasikristalle über einen langen Zeitraum nicht anerkannt wurde. Der Nobelpreisträger Linus Pauling wies seine Arbeit mit der Aussage „There are no quasi-crystals, there are quasi-scientists" (Shechtman, 2021, S. 5) zurück, da sie sich nicht in die wissenschaftlichen Erkenntnisse der damaligen Zeit einordnen ließ (ebd.). Im Jahr 2011 erhielt Shechtman den Nobelpreis für Chemie für seine Entdeckung der Quasikristalle. Das Phänomen, neue Erkenntnisse trotz Evidenzen zunächst abzulehnen, wenn sie den eigenen Glaubenssätzen widersprechen, ist auch wissenschaftlich beschrieben und wird als Semmelweis-Reflex bezeichnet. Dieser Begriff geht auf den Mediziner Ignaz Semmelweis zurück, dessen Arbeit und Erkenntnisse aus dem Jahr 1847 über den Zusammenhang von Hygiene und Kindbettfieber zunächst abgelehnt wurden und erst nach seinem Tod Anerkennung fanden (Gupta et al., 2020). Zahlreiche weitere Beispiele aus der Wissenschaft illustrieren diesen Effekt, etwa die Arbeiten Mendels zur Vererbungslehre, Pasteurs zu Krankheitserregern oder in jüngerer Zeit Allisons, welcher das Immunsystem in die Behandlung von Krebskrankheiten miteinbezog (ebd.).

Die Thematisierung von NOS trägt zwar bereits zur Sensibilisierung für die soziokulturelle Eingebundenheit, Kreativität und Subjektivität bei und fördert somit ein umfassenderes Verständnis der Wechselwirkungen zwischen Menschen und Naturwissenschaften, geht den Weg zu den dahinterliegenden Denkprozessen jedoch nicht ganz zu Ende. Die tiefergehende Frage wäre, warum die Gesellschaft einen so starken Einfluss auf die Naturwissenschaften hat oder was der Objektivität der Forschenden im Weg steht. Diese Frage ist nur mit der Art des menschlichen Denkens selbst zu beantworten, weshalb auch und gerade in den Naturwissenschaften ein tieferes Verständnis der zugrundeliegenden Denk- und Handlungsprozesse einen gewinnbringenden Schritt darstellen würde. Ein Verständnis von NOS zeigt sich damit als wichtig und relevant, ist aber ebenso wie das Fachwissen nur eine notwendige, jedoch keine hinreichende Bedingung für Kritisches Denken. NOS fokussiert sich zumeist auf die Wechselwirkungen zwischen dem Menschen und dem Prozess der Genese wissenschaftlicher Erkenntnisse, jedoch unterliegt auch die Rezeption naturwissenschaftlicher Ergebnisse kognitiven Verzerrungen, welche die Wissenschaftskommunikation erschweren können.

Zusätzlich zum Fachwissen und NOS ist es daher auch und gerade in den Naturwissenschaften von Bedeutung, für das eigene Denken zu sensibilisieren und die Fähigkeit zur Metakognition zu fördern, um kognitive Verzerrungen bewusst zu machen und ihre Auswirkungen zu reduzieren (Gigerenzer, 1991; Greenhalgh, 2002; Maddux, 1995; Taylor et al., 2009).

3.3 Die Definition Kritischen Denkens

Die bestehenden Definitionen Kritischen Denkens (vgl. Kap. 2.2) lassen sich im Hinblick auf Gemeinsamkeiten und Unterschiede in verschiedene Schwerpunktcluster einteilen.

Der gängigste Definitionsansatz, insbesondere in den Bildungswissenschaften und Naturwissenschaftsdidaktiken, beschreibt Kritisches Denken als die Kombination kognitiver Fähigkeiten und Dispositionen (Byrne & Johnstone, 1987; Ennis, 1987; Facione, 1990; Greene & Yu, 2016; Gunn et al., 2008; Pithers & Soden, 2000; Sinatra et al., 2014; Vieira et al., 2011; Yacoubian, 2020). Hierbei liegt der Schwerpunkt häufig auf der Ausbildung rationaler Fertigkeiten, wie der Interpretation, Analyse oder Bewertung von Informationen. Dispositionen, welche als Neigung zur Anwendung dieser kognitiven Fähigkeiten verstanden werden können, verbleiben meist auf einer deskriptiven Ebene. Sie charakterisieren die Eigenschaften, die ein*e kritisch Denkende*r aufweisen sollte, wie Aufgeschlossenheit, Neugier oder Selbstvertrauen (Facione, 1990), ohne jedoch darzulegen, wie diese Eigenschaften entwickelt oder gefördert werden können.

Vor allem psychologische und philosophische Ansätze stellen indes die Metakognition, die Fähigkeit, über sich selbst und seine kognitiven Prozesse nachzudenken, in das Zentrum ihrer Definition von Kritischem Denken (Dwyer et al., 2014; Halpern, 1999; D. Kuhn, 1999; Lipman, 1987; Olson & Astington, 1993; Paul & Elder, 2014). Der Fokus liegt damit auf dem Denkprozess selbst, der nicht nur als Fähigkeit, sondern als übergeordnetes Konstrukt betrachtet wird. Kritisches Denken fungiert in diesen Ansätzen als handlungsleitendes Konzept, das nicht zuletzt dazu dient, etwaigen Denkfehlern, wie kognitiven Verzerrungen, entgegenzuwirken und diese zu bewältigen (Gigerenzer, 1991; Greenhalgh, 2002; Maddux, 1995; Taylor et al., 2009).

Die affektive Dimension wird in den Definitionen des Kritischen Denkens nur vereinzelt thematisiert. Zwar wird ihre Relevanz in einigen erweiterten Ausführungen angedeutet, jedoch wird sie explizit nur von wenigen Autor*innen behandelt (Barr &

Ben-Shakhar, 2018; Brambatti Guzzo & Dall'Alba, 2017; Halpern, 2014; Rafolt et al., 2019). Erkenntnisse aus der Kognitions- und Verhaltenspsychologie verdeutlichen jedoch die zentrale Rolle affektiver Faktoren, wie Emotionen, Haltungen, Interessen, Motivationen und sozialen Normen, für Handlungs- und Entscheidungsprozesse, was bereits in den Wirtschaftswissenschaften Berücksichtigung findet (De Sousa, 2008; Elster, 1989; Festinger, 1957; Kahneman, 2012; Slovic et al., 2002; Thompson et al., 2013).

Diese unterschiedlichen Definitionsansätze verdeutlichen die Multidimensionalität Kritischen Denkens, der bei einer Neudefinition angemessen Rechnung getragen werden muss. Infolgedessen werden an dieser Stelle die Rahmenbedingungen zusammengefasst, die bei einer Neudefinition Kritischen Denkens handlungsleitend sind. Die Definition sollte:

- verschiedene Definitionsansätze berücksichtigen und diese zusammenführen,
- sowohl die kognitive als auch affektive Ebene einbeziehen,
- die metakognitive Komponente Kritischen Denkens berücksichtigen und
- der Komplexität des Konstrukts Rechnung tragen, aber gleichzeitig die Gesamtkomplexität so weit reduzieren, dass eine Operationalisierung möglich wird.

Insgesamt verdeutlichen diese Rahmenbedingungen, dass eine tragfähige Neudefinition des Kritischen Denkens die Mehrdimensionalität des Konstrukts auch in seiner Definition widerspiegeln und sowohl die kognitive und affektive Dimension als auch metakognitive Fähigkeiten einbeziehen muss. Nur so kann eine Grundlage geschaffen werden, die eine praktische Anwendung im Bildungskontext ermöglicht und gleichzeitig der Komplexität des Konstrukts gerecht wird.

Als Ausgangspunkt zur Modellierung einer Neudefinition, welche diesen Rahmenbedingungen gerecht wird, dient das *Synergiemodell des kritischen Denkens* von Rafolt et al. (2019), welches ein aktuelles und explizites Modell Kritischen Denkens darstellt (vgl. Kap. 2.2.1, Abbildung 8).

Das Modell zeigt das Involviertsein der denkenden Person mit einem Objekt oder Subjekt, begleitet von Faktoren, wie Wissen, Fähigkeiten, aber auch Haltung und Motivation sowie Normen, Werten und Emotionen. Unter Berücksichtigung der Erkenntnisse aus der Dual-Process Theory (vgl. Kap. 2.3.2, De Neys, 2006b; Frankish, 2010; Haidt, 2008; Kahneman, 2012) lässt sich dabei eine Unterscheidung in rationale und affektive Aspekte vornehmen, wobei die rationalen Elemente, wie Wissen und

Fähigkeiten, dem langsameren System 2 und die affektiven Elemente, wie Haltung, Motivation, Werte, Normen und Emotionen, dem intuitiven System 1 zugeordnet werden können.

Diese Differenzierung ist jedoch nicht nur hier anwendbar, sondern findet sich auch in den Definitionen verschiedener Begriffe der Naturwissenschaften und Naturwissenschaftsdidaktik wieder, wie bei der Kompetenzdefinition nach Weinert (2001). Hier lassen sich die „*kognitiven Fähigkeiten* und *Fertigkeiten*" (ebd., S. 27) dem rationalen System 2 und die „*motivationalen, volitionalen* und *sozialen Bereitschaften*" (ebd., S. 27 f.) dem intuitiven System 1 zuordnen (vgl. Kap. 2.1.1).

Die Bedeutung des affektiven Systems 1 wird zudem besonders deutlich bei der Betrachtung von NOS, wo Anteile, wie Subjektivität, Kreativität und soziokulturelle Eingebundenheit (Lederman, 1992), neben rationalen Aspekten, wie der empirischen Natur oder der wissenschaftlichen Methodik (Osborne et al., 2003), eine zentrale Rolle spielen. Auch das moralische Argument zur Vermittlung von NOS (Driver et al., 1996; Thomas & Durant, 1987) hebt moralische Werte und Normen hervor, die eng mit Prozessen des affektiven Systems 1 verknüpft sind (vgl. Kap. 2.1.2). Die NOSIS-Erweiterung (Höttecke & Allchin, 2020) (Abbildung 2) verweist neben den rationalen Faktoren ebenfalls auf explizit affektive Aspekte in Form der kognitiven und sozialen Filter bei der Rezeption naturwissenschaftlicher Erkenntnisse (vgl. Kap. 2.1.3).

Zu guter Letzt kann auch bei der Thematisierung von SSI ein starker Einfluss der affektiven Ebene und damit des Systems 1 festgestellt werden, da persönliche Werte und Glaubenssätze stark die Bewertung dieser Kontexte beeinflussen. Besonders bei moralisch komplexen Problemstellungen führt dies häufig zu einer Vermeidung einer klaren Positionierung, weshalb Lee und Yang (2019) explizit eine stärkere Berücksichtigung moralischer und psychologischer Elemente im Umgang von SSI fordern (vgl. Kap. 2.1.6, Cebesoy & Oztekin, 2018; Lee & Yang, 2019).

In Anlehnung an die Dual-Process Theory kann das *Synergiemodell des kritischen Denkens* von Rafolt et al. (2019) daher um die Konzepte dieser Theorie erweitert werden, was zur Entwicklung des Dual-Process Model of Critical Thinking (DP-CT) führt (Abbildung 11).

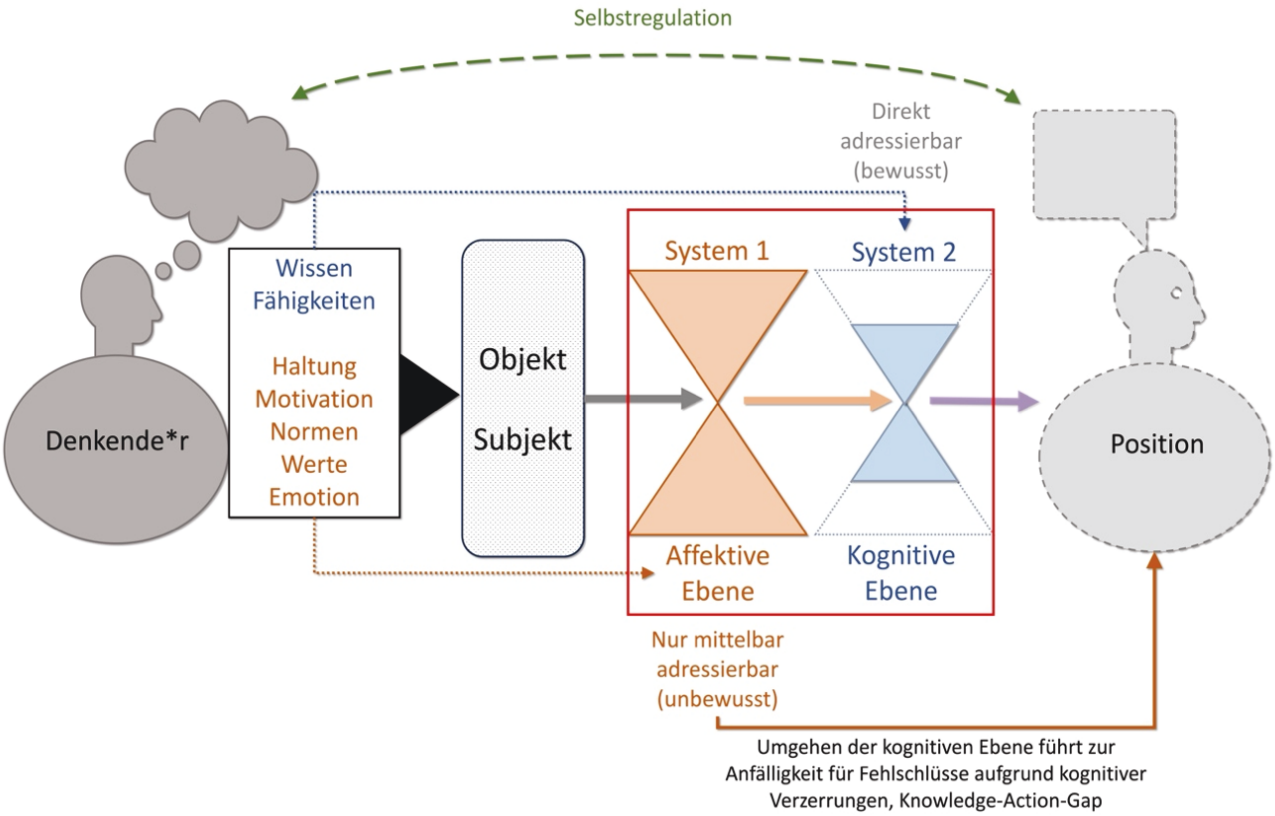


Abbildung 11: Dual-Process Model of Critical Thinking (DP-CT) (eigenes Modell)

Das *Dual-Process Model of Critical Thinking (DP-CT)* versteht Kritisches Denken als Positionierungsprozess, der bei der Auseinandersetzung mit einem Subjekt oder Objekt von kognitiven und affektiven Faktoren begleitet wird, welche die Wahrnehmung und Verarbeitung von Informationen innerhalb dieses Prozesses steuern. Das intuitive, affektive System 1 ist dem kognitiven, rationalen System 2 dabei vorgelagert, was zu Fehlschlüssen führen kann, wenn das rationale System nicht als Kontrollinstanz greift. In diesem Kontext bedeutet Positionierung nicht notwendigerweise einen aktiven Bewertungsprozess, sondern bezeichnet vielmehr den Weg der Informationsverarbeitung, welcher zur Wahrnehmung der Welt und zur Positionierung in ihr führt. Die/der kritisch Denkende ist in diesem Prozess bewusst involviert und kann beide Systeme zueinander in Bezug setzen. Dieses Bewusstsein ermöglicht eine kontinuierliche Selbstregulation, Neubewertung und Anpassung an Informationen, unter Anwendung der ihr*m zur Verfügung stehenden kognitiven Fähigkeiten und Berücksichtigung affektiver Einflüsse.

Basierend auf dem hier vorgestellten *Dual-Process Model of Critical Thinking (DP-CT)* wird an dieser Stelle folgende Neudefinition Kritischen Denkens vorgenommen:

Kritisches Denken (KD) bezeichnet eine sorgfältige, systematische und vor allem begründete Entscheidungsfindung bei der Auseinandersetzung der*s Denkenden mit einem Objekt oder Subjekt. Diese Auseinandersetzung beinhaltet die Evaluation und Anwendung des eigenen Wissens und der eigenen Fähigkeiten auf kognitiver Ebene sowie die Berücksichtigung der eigenen Emotionen, Werte, Einstellungen und Motivationen auf affektiver Ebene. Es umfasst zudem eine konstante Selbstreflexion und -regulation, die eine Rückkopplung zur*m Denkenden selbst ermöglicht und diese*n nicht nur in die Position der*s Bewertenden, sondern auch in die Position der*s zu Bewertenden versetzt (eigene Definition u. a. in Anlehnung an Haidt, 2008; Halpern, 2014; Kahneman, 2012; D. Kuhn, 1999; Rafolt et al., 2019).

Auf dieser Definition begründet kommt Kritischem Denken eine Schlüsselrolle zu, da es nicht nur eine Fähigkeit, sondern den Prozess geleiteten Denkens selbst darstellt. Das Kritische im Denken ist somit sein multiperspektivisch-prüfender Charakter, welcher nicht allein die Anwendung des Wissens und der Fähigkeiten bestimmt, sondern auch berücksichtigt, in welchem affektiven Kontext diese zur Ausprägung kommen.

Ein besonderes Augenmerk liegt in diesem Kontext auf der metakognitiven Fähigkeit der Selbstreflexion, welche dem Menschen eigen ist und eine der am weitesten fortgeschrittenen Denkfähigkeiten darstellt (Tulving, 2005). Selbstreflexion wird in der Literatur häufig definiert als die Einsicht und Bewertung der eigenen Gedanken, Gefühle und des Verhaltens (Carver & Scheier, 2001; Grant et al., 2002; Silvia, 2021; Silvia & Phillips, 2011). Facione (1990) konkretisiert dies und spricht von Selbst-Überprüfung, unter die er folgende Punkte fasst:

- Das Reflektieren über das eigene Denken und die Überprüfung sowohl der erzielten Ergebnisse als auch der korrekten Anwendung und Ausführung der damit verbundenen kognitiven Fähigkeiten,

- das objektive und durchdachte metakognitive Bewerten der eigenen Meinungen sowie der Gründe, die zu deren Annahme führen,

- das Beurteilen, in welchem Maß das eigene Denken durch Wissenslücken, Stereotype, Vorurteile, Emotionen oder andere Faktoren beeinflusst wird, welche die Objektivität oder Rationalität einschränken und

– das Reflektieren über die eigenen Motivationen, Werte, Einstellungen und Interessen, um festzustellen, inwieweit man sich darum bemüht hat, unvoreingenommen, fair, gründlich, objektiv, wahrheitsgetreu, vernünftig und rational bei der Formulierung von Analysen, Interpretationen, Bewertungen, Schlussfolgerungen oder Äußerungen zu sein (vgl. ebd., S. 19).

In Summe kann Selbstreflexion im Kontext des Kritischen Denkens wie folgt definiert werden:

Selbstreflexion ist ein Prozess der *introspektiven Überprüfung* eigener Gedanken, Gefühle, Verhaltensweisen und Prägungen (Motivation, soziales Umfeld, eigene Erfahrungen u. v. m.) sowie darauf aufbauend ein *Bewusstsein für interne und externe Einflussfaktoren*. Sie umfasst die *Wahrnehmung, Gewichtung und Bewertung* ebenjener, um diese *in Entscheidungsprozessen berücksichtigen* zu können und so eine objektivere Position zu erhalten (eigene Definition in Anlehnung an Carver & Scheier, 2001; Facione, 1990; Gigerenzer, 1991; Grant et al., 2002; Silvia, 2021).

Auch die Charakterisierung der Selbstregulation erfolgt im Rahmen dieser Konzepte:

Selbstregulation bezeichnet die Möglichkeit und Bereitschaft zur selbstständigen Anpassung der eigenen Handlungen sowie Einflussnahme auf eigene Gedanken und Emotionen. Es umfasst das Überwachen des eigenen Denkens sowie das Bewerten der eigenen kognitiven Prozesse und affektiven Einflüsse. Damit stellt es im Gegensatz zur Selbstreflexion, welche primär eine selbstbezogene Beobachtung darstellt, die darauf aufbauende Selbstadaption dar (eigene Definition in Anlehnung an Facione, 1990; Halpern, 2014; Paul & Elder, 2014; Rafolt et al., 2019).

Kritisches Denken bildet in Summe das Fundament für das Wissen, die Fähigkeiten sowie die Selbstreflexion und definiert damit den Rahmen, in dem diese zur Anwendung kommen. Es kann als grundständige Denkpraxis bezeichnet werden, wohingegen das Wissen das Werkzeug ist, welches vom Denken kritisch geführt und angewandt wird. In Konsequenz ermöglicht Kritisches Denken die Anpassung und Selbstregulation eigener Entscheidungen und getroffener Bewertungen, nicht zuletzt auch sich selbst betreffend.

Kritisches Denken, wie es in dieser Arbeit definiert wird, weist einen starken systemischen Charakter auf, da keiner der beiden Dimensionen Kritischen Denkens isoliert betrachtet werden sollte (Abbildung 11). Die affektiven und kognitiven Dimen-

sionen bedingen einander und sind rein evolutiv nicht trennbar (vgl. Kap. 2.3.2). Insbesondere die affektive Ebene lässt sich nicht umgehen, sondern lediglich bewusst machen. Diese systemische Verflechtung bildet die Grundlage für eine *Rationalität* im Sinne Stanovichs (2008), die sowohl die algorithmische Ebene der grundlegenden kognitiven Prozesse als auch das intentionale Denken, welches selbstreflexives und selbstregulatives Denken umfasst, integriert (vgl. Kap. 2.3.4). Kritisches Denken fungiert dabei als Weg zu dieser Art Rationalität, indem es ein Abwägen zwischen beiden Dimensionen ermöglicht, und kann als Brücke zwischen analytischen und holistischen Denkansätzen gesehen werden. Durch diese Betrachtungsweise verschiebt sich der Fokus von einer statischen zu einer dynamischen Sichtweise, welche die Grenzen des Systems überschreitet (Reiners, 1996). Durch seinen dynamischen Ansatz fördert Kritisches Denken nicht nur das Verständnis komplexer Zusammenhänge, sondern auch die Fähigkeit zur Anpassung an neue Erkenntnisse und veränderte Bedingungen. In diesem Rahmen stellt die isolierte analytische Betrachtung eines Systems „eine Art Momentaufnahme des klassifizierenden Ordnens dar“ (ebd., S. 13), während die systemische Denkweise als geeigneter Mittelweg für die Analyse komplexer Zusammenhänge zwischen Naturwissenschaften, Gesellschaft, Ökonomie und Ökologie erscheint. Kritisches Denken zielt daher nicht auf eine isolierte Betrachtung ab, sondern strebt ein umfassendes Verständnis des Gesamtsystems und seiner Wechselwirkungen darüber hinaus an.

Wissen ist kontextgebunden, Einflüsse sind universell und das Kritische Denken ist die transzendierende Kraft, welche diese disparat scheinenden Elemente verbindet und ihnen durch Reflexion Kohärenz verleiht, wodurch eine neue Ebene der Erkenntnis erreicht werden kann. In der Konsequenz kann Kritisches Denken nicht nur als systemisch, sondern sogar als emergent betrachtet werden, da in diesem hierarchisch übergeordneten Niveau neue qualitative Eigenschaften entstehen (Primas, 1985). Ähnlich einem Aluminium-Cluster, welches seine metallischen Eigenschaften erst bei einer bestimmten Größe erhält (Cândido et al., 2012; Rottener et al., 2024), erlangt Kritisches Denken erst im Zusammenspiel seiner Bestandteile, der affektiven und kognitiven Ebene sowie dem Fachwissen, epistemologischen Verständnis und der Selbstreflexion, seine übergreifend reflexiven Eigenschaften und befähigt zur umfassenden Urteilsbildung.

3.4 Abgrenzung zu anderen Konzepten

Kritisches Denken muss an dieser Stelle klar von anderen Konzepten, wie dem *Systems Thinking*, abgegrenzt werden. Systems Thinking bezieht sich auf eine methodische Herangehensweise zur Analyse und Vorhersage des Verhaltens komplexer Systeme. Der Fokus liegt auf dem Erkennen von Beziehungen zwischen Systemelementen, dem Verstehen von Rückkopplungsschleifen und der Identifikation von Variablen, die die Dynamik des Systems beeinflussen (Shaked & Schechter, 2017). Um die Komplexität solcher Systeme zu bewältigen, erfolgt häufig eine Reduktion durch konzeptionelle Modellierung (Arnold & Wade, 2015). Dabei liegt der Schwerpunkt auf der Analyse von objektiven Prozessen und Strukturen, weshalb Systems Thinking vorwiegend in organisatorischen Systemen angewendet wird (Shaked & Schechter, 2017).

Im Gegensatz dazu zielt Kritisches Denken nicht nur auf die Analyse von externen Strukturen ab, sondern integriert sowohl kognitive als auch affektive Einflüsse explizit in den Denkprozess. Kritisches Denken erweitert die systemische Perspektive, indem es die Reflexion über persönliche Einstellungen, Emotionen und Werte einbezieht, was es auch für individuelle Entscheidungsprozesse relevant macht. Während Systems Thinking primär auf die Modellierung und Vorhersage von Systemverhalten abzielt (Arnold & Wade, 2015), fokussiert Kritisches Denken auf den metakognitiven Prozess der Selbstreflexion und Selbstregulation, um Verzerrungen zu erkennen und fundierte, rationale Entscheidungen zu treffen.

Systems Thinking konzentriert sich zudem stark auf explizites Wissen über systemische Strukturen und Prozesse, das durch Modelle und Analysen greifbar gemacht wird. Es setzt auf formale Werkzeuge, um das Verhalten von Systemen zu verstehen und gezielte Anpassungen vorzunehmen (Arnold & Wade, 2015; Monat & Gannon, 2015). Kritisches Denken hingegen geht über die reine Anwendung von explizitem Wissen hinaus und bezieht auch implizites, intuitives Wissen ein, das aus affektiven Reaktionen und unbewussten Prozessen resultiert. Durch die bewusste Reflexion der affektiven Ebene ermöglicht Kritisches Denken die Identifikation kognitiver Verzerrungen und eine ausgewogene Entscheidungsfindung. Es umfasst nicht nur den rationalen Umgang mit Wissen, sondern auch die Fähigkeit, den eigenen Denkprozess und die damit verbundenen affektiven Einflüsse kontinuierlich zu reflektieren.

Ein weiteres Konzept, welches an dieser Stelle vom Kritischen Denken abgegrenzt werden soll, ist die Bewertungskompetenz. Diese ist in den aktuellen weiterentwickelten Bildungsstandards definiert als Ausdruck der

> Kenntnis von fachlichen und überfachlichen Perspektiven und Bewertungsverfahren verbunden mit der Fähigkeit, Handlungsoptionen anhand verschiedener Kriterien zu beurteilen, um Entscheidungen auch auf ethischer Grundlage zu treffen, die Folgen abzuschätzen und Entscheidungsprozesse zu reflektieren (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024b, S. 7, 2020, S. 20).

Die Bewertungskompetenz ist eine der vier explizit benannten Kompetenzbereiche der Bildungsstandards und damit ein wesentlicher Bestandteil des naturwissenschaftlichen Unterrichts. Überdies beinhaltet sie die Reflexion über den Wert und die Grenzen naturwissenschaftlicher Erkenntnisse und geht damit über das reine Verständnis von Inhalten hinaus (Sander & Höttecke, 2017).

Obwohl sich Bewertungskompetenz und Kritisches Denken in einigen Aspekten überschneiden, gibt es wesentliche Unterschiede. Die Bewertungskompetenz ist stark auf die Anwendung von Wissen in spezifischen Kontexten fokussiert und bietet schwerpunktmäßig einen Rahmen für den Einsatz von Fachwissen und kognitiven Fähigkeiten zur Bewertung von Informationen. Dies zeigt sich nochmal verstärkt im Beitrag zur Implementation der Bildungsstandards für Lehrkräfte, welcher explizit darauf hinweist, dass die einzelnen Standards der Bewertungskompetenz „dahingehend ausgeschärft [wurden], dass sie nun den Prozess der Bewertung von Sachverhalten und Handlungsoptionen aus fachlicher Perspektive stärker Rechnung tragen“ (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024a, S. 14). Bei diesem stärker inhaltszentrierten Ansatz stellt sich allerdings die Frage, ob im Hinblick auf den Kompetenzbegriff davon ausgegangen werden kann, dass Handeln allein die Anwendung von Wissen ist. Mit Blick auf die Definition des Verfügungswissens von Mohr (1989), welches die Antwort darauf zu geben versucht, wie etwas getan werden kann (Reiners, 2022), stellt eine stärkere fachliche Perspektive eher eine Instruktion dafür dar, welcher Gestalt eine Handlung sein kann, als die tatsächliche Befähigung zur selbstständigen Entscheidung, welche Handlung wann angebracht ist. Im Gegensatz dazu enthält richtiges Orientierungswissen eine starke moralisch-ethische Komponente zur Beantwortung der Fragen „Was soll ich tun?“ oder „Was darf ich […] tun?“ (Mohr, 1989, S. 200; Reiners, 2022, S. 96), weswegen Kompetenz nicht direkt als Orientierungswissen verstanden werden kann.

Kritisches Denken hingegen hat einen breiteren Fokus, welcher auch ethisch-moralische Aspekte einschließt, und kann als übergeordnetes Konstrukt verstanden werden, in dem die Bewertungskompetenz als eine Ausprägung Kritischen Denkens in Erscheinung treten kann. Es beruht dabei nicht nur auf naturwissenschaftlichem Fachwissen und der Berücksichtigung gesellschaftlicher Normen, sondern umfasst auch die metakognitive Komponente der Selbstreflexion, welche in die Lage versetzt, nicht nur Informationen zu analysieren, sondern auch die eigenen Denkprozesse in einen Bezug zu ihnen zu setzen. Kritisches Denken ist somit umfassender und systemisch orientiert, indem die kognitive und affektive Dimension miteinander wechselwirkend betrachtet werden. Es geht nicht nur darum, Urteile zu fällen, sondern auch darum, den gesamten kognitiven Prozess, einschließlich der affektiven Einflüsse, zu reflektieren und zu regulieren, um ausgewogene und rationale Entscheidungen treffen zu können (vgl. Kap. 3.3).

Somit stellt Kritisches Denken eine Erweiterung der primär fachimmanenten Kompetenzbereiche dar, indem nicht ausschließlich das Wissen und dessen kontextualisierte Anwendung im Fokus stehen, sondern der Umgang mit dem Denken und dem Wissen selbst in den Mittelpunkt der Betrachtung rückt. Damit verfolgt Kritisches Denken das Ziel, den Bildungsprozess über das Fachwissen hinaus zu gestalten und mittels des naturwissenschaftlichen Unterrichts „Education through Science“ (Holbrook, 2010, S. 87) zu ermöglichen, um zur gesellschaftlichen Teilhabe und zur lebensweltlichen Orientierung beizutragen.

3.5 Kritisches Denken im naturwissenschaftlichen Unterricht

Die Anerkennung wissenschaftlicher Werte kann nur mit einem reflektierten Verständnis der Wissenschaft einhergehen. Eine rein skeptische Haltung ist dabei nicht ausreichend, um ihre Stärken und Grenzen zu reflektieren, und bietet nur wenig praktischen Nutzen in alltäglichen Entscheidungssituationen (Collins, 2009; Sharon & Baram-Tsabari, 2020). Die multiperspektivische Denkpraxis Kritischen Denkens im Sinne der vorgestellten Definition und die Berücksichtigung sowohl kognitiver als auch affektiver Aspekte beinhaltet hingegen das Potenzial, komplexe Zusammenhänge innerhalb der Naturwissenschaften sowie deren Interdependenzen mit externen Faktoren, wie gesellschaftlichen, ethischen und ökologischen Aspekten, adäquat zu erfassen. Dabei dient das Kritische Denken als Vermittlerinstanz zwischen der affektiven und kognitiven Dimension, welche gleichzeitig in dessen Fundament eingebettet sind (Abbildung 12).

Rationale Aspekte, wie die naturwissenschaftliche Grundbildung, das epistemologische Wissen rund um NOS sowie die Erweiterung um die mediale Grundbildung und NOSIS, ermöglichen dabei ein Aneignen von Grundwissen, die Entwicklung eines Verständnisses der naturwissenschaftlichen Erkenntnisgewinnung, den Erwerb von Argumentationsfähigkeiten, die Bewertung, Analyse und Einordnung von Informationen sowie ein Bewusstsein für die Faktoren, welche eine gesellschaftliche Aufnahme naturwissenschaftlicher Informationen begleiten. Die affektive Dimension hingegen ermöglicht die Selbstreflexion *interner Faktoren*, wie Motive, Normen, Werte, Emotionen, Vorurteile oder Überzeugungen, sowie die Selbstreflexion *externer Faktoren*, wie Medien, soziale Normen und Kontexte. Hinzu kommt die Selbstregulation, die Anpassung des eigenen Verhaltens basierend auf den gemachten introspektiven Beobachtungen.

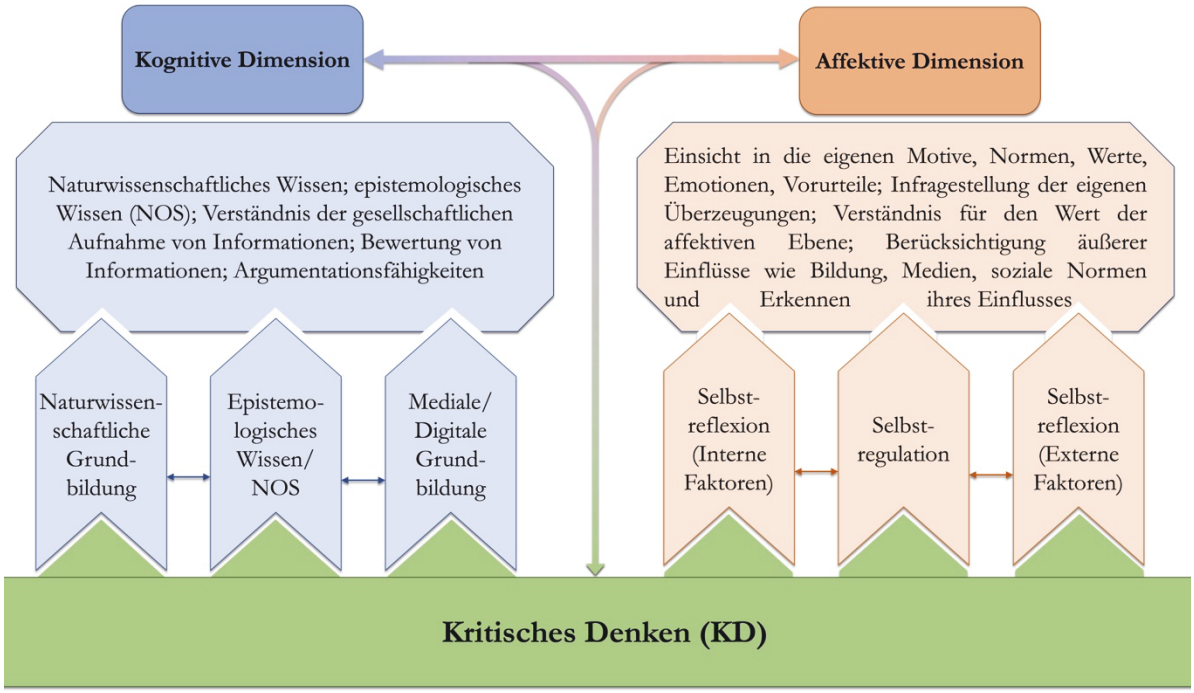


Abbildung 12: Kritisches Denken als Fundament und Vermittlerinstanz zwischen kognitiver und affektiver Dimension (eigene Darstellung)

Durch das Bewusstwerden von unbewussten Faktoren, wie Emotionen, Vorurteilen oder sozialen Einflüssen, wird der/die kritisch Denkende in die Lage versetzt, diese gegen vorhandenes Wissen und rationale Überlegungen abzuwägen. Dabei ist der/die kritisch Denkende in erster Linie nicht dadurch charakterisiert, welche Entscheidung getroffen wird, sondern wie sich einer Problemstellung, Information oder Herausforderung genähert wird (Brambatti Guzzo & Dall'Alba, 2017).

Bezogen auf das Bildungsziel, Wissen in Handeln umsetzbar zu machen, was sich neben der Kompetenzdefinition (Weinert, 2001) auch in der Definition der naturwissenschaftlichen Grundbildung zeigt (Organisation for Economic Co-operation and Development, 1999) (vgl. Kap. 2.1.1), sind nicht nur Wissen und Fähigkeiten relevant, sondern auch und vor allem psychosoziale Ressourcen (Rychen und Salganik, 2001; vgl. Kap. 2.1.1). Kritisches Denken bietet die Möglichkeit, zwischen beiden Ebenen zu moderieren und intuitive Einflüsse des Systems 1 auf die bewusste Ebene zu heben, um sie im Entscheidungsprozess zu reflektieren. Dabei ist es nicht das Ziel, die affektive Ebene zu unterdrücken, sondern diese bewusst in Entscheidungs- und Handlungsprozesse einzubeziehen.

Reine fachwissenschaftliche Vermittlung kann in der Lebensrealität der Lernenden dazu führen, dass dort Anwendungsbereiche des Wissens nicht erkannt und in Konsequenz Wissen und Fähigkeiten nicht eingesetzt werden. Die Implementation von Nature of Science, besonders unter Berücksichtigung des demokratischen Arguments zur Förderung einer gesellschaftlichen Mündigkeit (vgl. Kap. 2.1.2), gewinnt damit an Bedeutung, da sich in einer immer stärker naturwissenschaftlich und technisch geprägten Welt ein konkreter Lebensweltbezug für Lernende zeigt. Um aber das Wesen der Naturwissenschaften und die Position des Menschen in diesem Kontext zu verstehen und auch resistenten Mythen bezüglich NOS zu begegnen (Mesci & Schwartz, 2017; S. Müller, 2021; S. Müller & Reiners, 2021), wohnt dem Kritischen Denken das Potenzial inne, für die Verflechtungen zwischen Mensch und Naturwissenschaften zu sensibilisieren und die tieferliegenden, damit verbundenen affektiven Prozesse zu verdeutlichen, die letztlich zur Subjektivität und soziokulturellen Eingebundenheit führen. Dabei sollte die Naturwissenschaftsdidaktik keine Bedenken haben, die Naturwissenschaften damit abzuwerten, indem das Bild der objektiven und unbeeinflussbaren Wissenschaftler*innen relativiert wird, sondern es sollte eine Unterscheidung geschehen zwischen den Ansprüchen, Mechanismen und Methoden der Wissenschaft und denen, die sie ausüben. Dies könnte in der Konsequenz sogar die Wissenschaftskommunikation verbessern, indem der Mensch, einschließlich seiner Einflüsse und Verzerrungen, losgelöst vom Anspruch der Wissenschaft selbst betrachtet werden kann und die Attributionen der Wissenschaft nicht unkritisch auf die Ausübenden übertragen werden.

Kritisches Denken spielt im Weiteren auch eine große Rolle bei der individuellen Rezeption von Informationen und der Einordnung ihrer Wissenschaftlichkeit, was nicht zuletzt aufgrund des Einflusses der neuen Medien an Bedeutung gewinnt. Dies schlägt sich auch in den weiterentwickelten Bildungsstandards nieder, die seit 2024

den Umgang mit digitalen Medien explizit als Bestandteil naturwissenschaftlicher Bildung benennen (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024b). Durch die sozialen Medien und die damit einhergehende Freiheit der ungeprüften Informationsverbreitung (vgl. Kap. 2.1.3) bedarf es eines reflektierteren Umgangs mit Informationen und damit verbundenen Verzerrungen, um in diesem Kontext wahrlich kritisch an die Bewertung heranzutreten. Auch die Abgrenzung wissenschaftlicher Arbeiten von pseudowissenschaftlichen Ansätzen profitiert von einer Berücksichtigung der affektiven Ebene, besonders, da die Abgrenzungskriterien nicht eindeutig trennscharf sind (vgl. Kap. 2.1.5). Infolgedessen ist eine multiperspektivische Betrachtung derlei Informationen unter Berücksichtigung des eigenen Wissens, des Wissens um den Erkenntnisgewinnungsprozess und der Selbstreflexion unabdingbar, um zur Rationalität beizutragen. Kritisches Denken kann auf diese Weise dazu beitragen, pseudowissenschaftliche oder alternativmedizinische Behauptungen zu prüfen und für sich selbst informiertere und bewusstere Entscheidungen zu treffen.

Im Hinblick auf die digitale Bildung verdient auch der Vormarsch der *Künstlichen Intelligenzen* (KI) eine gesonderte Betrachtung. Der Umgang mit KI scheint auf den ersten Blick einfach und effektiv zu sein, birgt aber nicht direkt absehbare Risiken. Auch große Sprachmodelle (Larg Language Model, LLM), wie ChatGPT, können Fragen falsch verstehen, Informationen verzerren, Fake News verbreiten oder diskriminieren. Zudem sind besonders vortrainierte Modelle, sogenannte Generative Pre-Trained Transformer (GPT), stark abhängig von den Informationen, mit denen sie trainiert wurden. Infolgedessen ist es relevant, die von der KI gegebenen Informationen zu prüfen und nicht unreflektiert zu übernehmen. Dies setzt die Fähigkeit zum Kritischen Denken voraus, um die Informationen mit dem eigenen Wissen und Standpunkten abzugleichen (Bucher et al., 2024).

Zudem vollzieht sich bei der Nutzung von Künstlicher Intelligenz ein Rollenwechsel, der den Übergang von der aktiven, selbstständigen Bearbeitung von Aufgaben hin zur Steuerung und Überwachung der KI-basierten Prozesse markiert. Anstelle des eigenständigen Handelns tritt zunehmend die Verantwortung für das Management der Künstlichen Intelligenz. Eine Künstliche Intelligenz ist damit lediglich so leistungsfähig wie die/der Nutzende, denn ihre Effektivität hängt maßgeblich von der Kompetenz, den Entscheidungen und der sorgfältigen Eingabe derjenigen ab, die sie steuern. In Konsequenz ist die Qualität der durch die KI erzeugten Ergebnisse direkt an die Fähigkeit der*s Nutzenden gebunden, relevante Daten bereitzustellen,

zielgerichtete Anweisungen zu geben und die Resultate kritisch zu prüfen sowie anzupassen. Dieser Wandel erfordert eine Neuausrichtung der Kompetenzprofile, wobei insbesondere Kritisches Denken und die Fähigkeit zur Selbstreflexion im Umgang mit KI unverzichtbar ist (ebd.).

Einen besonderen Stellenwert erhält Kritisches Denken zudem in der Thematisierung von *Socioscientific Issues*, welche wegen ihrer kontroversen und komplexen Natur einen ebenso mehrdimensionalen Bewertungsansatz erfordern. Aufgrund der Verflechtungen der Naturwissenschaften mit gesellschaftlichen, wirtschaftlichen, ökologischen und ethischen Faktoren, ist die Befähigung zu multiperspektivischem Denken, welches in der Lage ist, diese Aspekte in einen Bezug zueinander zu setzen, von besonderer Relevanz (vgl. Kap. 2.1.6). Auch die Uneindeutigkeit innerhalb von SSI stellt eine große Herausforderung dar, besonders wenn sich diese moralisch fordernd zeigen. Als Reaktion versuchen Lernende in diesen Kontexten oft eine klare Stellungnahme zu vermeiden (Chen & Xiao, 2021; Ideland et al., 2011). Diesem Konflikt könnte mit einer Sensibilisierung für die Vorläufigkeit naturwissenschaftlicher Erkenntnisse und Förderung der Selbstreflexion begegnet werden, indem zum einen auf den wissenschaftlichen Prozess selbst rekurriert und zum anderen die Bewertung nicht als Teil der persönlichen Identifikation wahrgenommen wird. Durch das Begreifen von Bewertungsprozessen als wohlbegründete Entscheidung auf Basis der aktuellen Datenlage und nicht als Teil der Persönlichkeit, kann sowohl die Entscheidungsschwelle gesenkt als auch mit kognitiven Verzerrungen, wie dem Commitment Bias, dem Festhalten an bereits getroffenen Entscheidungen, umgegangen werden (Staw, 1976). Durch den höheren Stellenwert der Selbstreflexion bei der Bewertung von SSI wird auch der Forderung Lee und Yangs (2019) entsprochen, bei der Thematisierung von SSI im Unterricht psychologische Faktoren verstärkt zu berücksichtigen, um einen Raum zur Reflexion eigener Werte und Überzeugungen zu schaffen (vgl. Kap. 2.1.6).

Dies zeigt auf, dass naturwissenschaftliches Wissen sowie ein Verständnis der Epistemologie durch NOS für Kritisches Denken bei der Bewertung von SSI zwar notwendige, jedoch keine hinreichenden Bedingungen darstellen (Ennis, 1989). Nur ergänzt durch die Selbstreflexion interner und externer Faktoren und ein Bewusstsein für deren Einflüsse können voreilige Schlüsse überdacht und revidiert, bestehende Ansichten hinterfragt und soziale Kontexte miteinbezogen werden (Lee & Yang, 2019).

Resultierend lassen sich drei Schwerpunkte Kritischen Denkens im naturwissenschaftlichen Unterricht ableiten, welche sich ebenfalls systemisch bedingen (Abbildung 13):

- Das naturwissenschaftliche Fachwissen stellt eine fundierte Grundlage dar, auf der basierend begründete Bewertungen und Handlungen getroffen werden können.

- Das epistemologische Verständnis, gestützt auf die Konzepte von NOS und NOSIS, befähigt dazu, die Verlässlichkeit und Wissenschaftlichkeit von Informationen einzuschätzen, angewandte Methoden korrekt einzuordnen und die Entstehung wissenschaftlicher Erkenntnisse nachzuvollziehen.

- Die Selbstreflexion ermöglicht es, interne und externe Einflussfaktoren wahrzunehmen und auf dieser Grundlage sich selbst zu regulieren, um kognitive Verzerrungen zu erkennen, den eigenen Standpunkt zu hinterfragen und gegebenenfalls neue Bewertungen vorzunehmen.

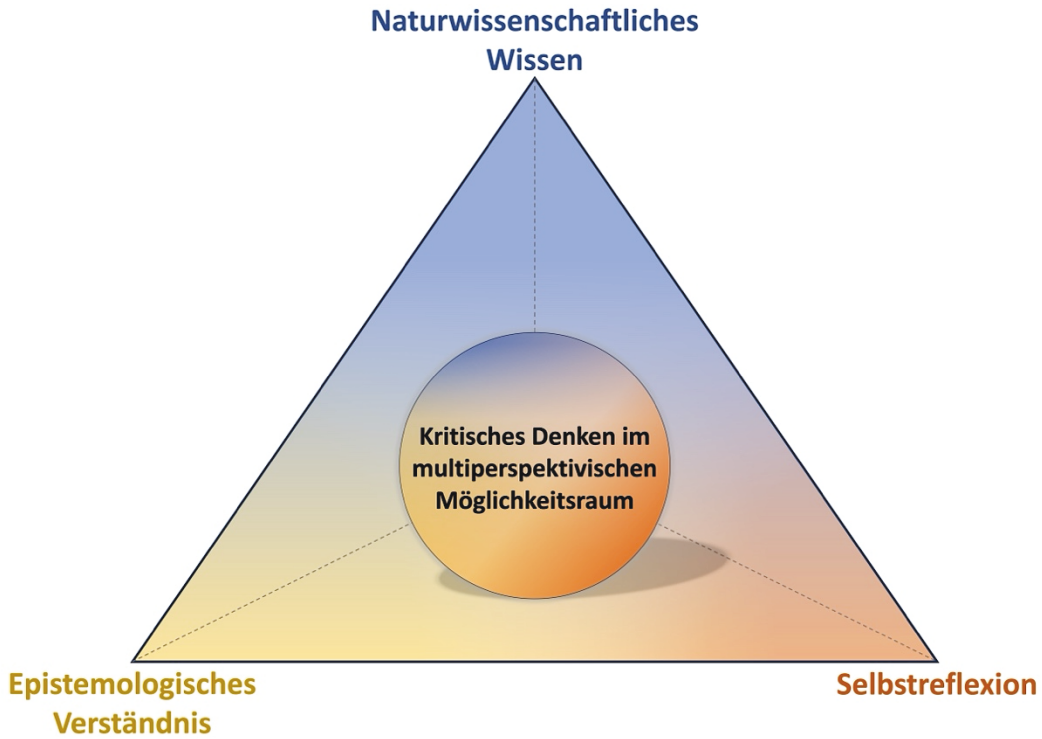


Abbildung 13: Triadisches Modell Kritischen Denkens in der Naturwissenschaftsdidaktik (eigenes Modell)

Daraus erwächst ein multiperspektivischer Möglichkeitsraum, der Lernende befähigt, naturwissenschaftliches Wissen im übergeordneten Kontext wahrzunehmen und zu bewerten. Es soll an dieser Stelle aber hervorgehoben werden, dass es nicht

darum geht, eine richtige Entscheidung zu vermitteln, sondern einen Möglichkeitsraum zu schaffen, in dem eigene reflektierte Entscheidungen gefällt werden können.

Ziel ist es daher, diese anderen Entscheidungsansätze so handhabbar zu machen, dass sie zur handlungsrelevanten Lösungsfindung einen subjektiv erfahrbaren Mehrwert im Gegensatz zum eigenen intuitiven (Bauch-)Gefühl bieten. Hier lassen sich Parallelen zu den Alltagsvorstellungen von SuS erkennen, da diese ebenso nicht auslöschbar sind und im Umgang mit dieser Tatsache die verfügbaren alternativen Sichtweisen attraktiver gemacht werden müssen (Duit et al., 1981; Reiners, 2022).

Um in diesem Kontext dem Ziel der Naturwissenschaftsdidaktik gerecht zu werden, Wissen und Fähigkeiten zu vermitteln, die auch praktisch Anwendung finden können, müssen Möglichkeiten geschaffen werden, in denen Kritisches Denken bewusst angewandt und geschult werden kann (Sadler, 2011; Zeidler & Kahn, 2014). Hierfür bietet die Thematisierung von SSI einen Bezugsrahmen für die Anwendung und Förderung Kritischen Denkens in der Chemiedidaktik (Rafolt et al., 2019) und kommt damit gleichzeitig der Forderung nach, dass Denkfähigkeiten am besten zielgerichtet innerhalb bestimmter Domänen vermittelt werden (Perkins, 1987). Dabei sollte Kritisches Denken explizit vermittelt und thematisiert werden (Facione, 1990; Kolstø et al., 2006), da sich auch bei ähnlich komplexen Konstrukten, beispielsweise bei NOS, gezeigt hat, dass explizite Vermittlungsansätze gegenüber impliziten Ansätzen effektiver sind (Abd-El-Khalick et al., 1998; Höttecke & Silva, 2011; Khishfe & Abd-El-Khalick, 2002; Lederman & Lederman, 2014).

Insgesamt kommt dem Kritischen Denken sowohl im Lernprozess naturwissenschaftlicher Charakteristika und Methodiken als auch in der Bewertung von SSI eine besondere Bedeutung zu, in dem es als Bindeglied zwischen den naturwissenschaftlichen und menschenbezogenen Aspekten fungiert. Vor diesem Hintergrund lässt sich die Anwendbarkeit naturwissenschaftlichen Wissens als ein Kontinuum begreifen, das sich mit wachsender Kontextualisierung in Bezug auf den Menschen steigert (Abbildung 14).

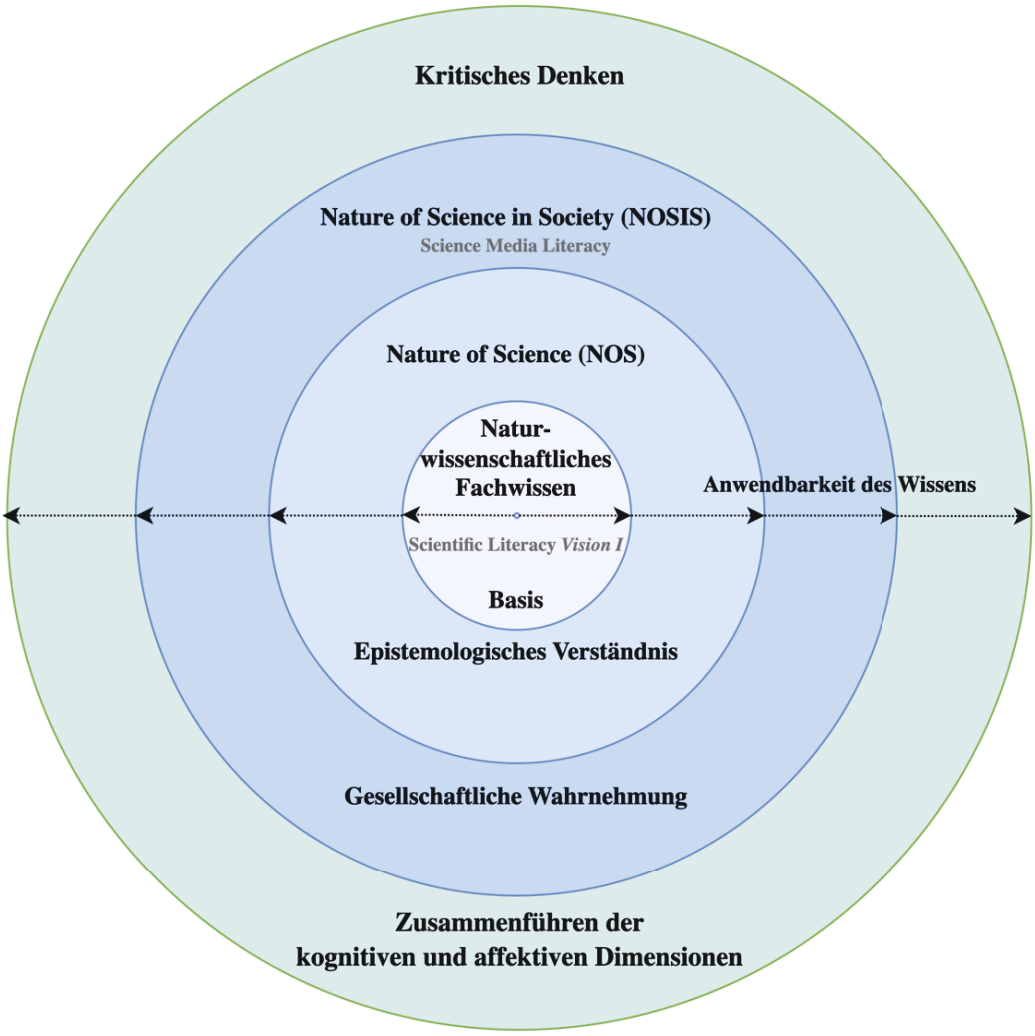


Abbildung 14: Ebenenmodell des Kritischen Denkens in der Naturwissenschaftsdidaktik (eigene Darstellung)

Während das reine naturwissenschaftliche Fachwissen im Sinne des Verfügungswissens (Mohr, 1989) noch eine geringe Anwendbarkeit in der persönlichen Lebensrealität der*s Einzelnen hat, baut ein Verständnis von NOS auf dem Fachwissen auf, ergänzt dessen Bedeutung in soziokulturellen Kontexten und hebt die Bedeutung des Menschen bei der Wissensgenese heraus (Reiners, 2022). NOSIS wiederum erweitert das Konzept um die Komponente der Digitalisierung inklusive der damit verbundenen Herausforderungen. Kritisches Denken hingegen verbindet alle Ebenen und umschließt sie als Grundlage, indem es die einzelnen kognitiven und affektiven Elemente systemisch miteinander verbindet und so eine bessere Abwägbarkeit sowie einen multiperspektivischen Möglichkeitsraum schafft, in dem auf Basis der vorangegangenen Faktoren individuelle Entscheidungen möglich sind. Damit wird das erworbene Fachwissen kontextualisiert und zum Orientierungswissen erhoben, dem ein tatsächlicher Nutzen zukommt, und so beispielsweise die gesellschaftliche

Partizipation bei naturwissenschaftlichen Fragestellungen und ein kontinuierliches Weiterlernen ermöglicht.

Überhaupt stellt es sich mittlerweile als unmöglich dar, sämtliches vorhandenes Fachwissen im Schulunterricht zu vermitteln. Das Wissen der Welt verdoppelt sich alle fünf bis zwölf Jahre (Lyman & Varian, 2003) und ist zudem fast vollständig digital abrufbar. In diesem Verlauf kann es somit nicht mehr das Ziel sein, eine große Quantität an Fachwissen zu vermitteln, sondern anhand exemplarischen Fachwissens ein qualitativ umfassenderes Verständnis im Umgang mit und der Aneignung von neuem Fachwissen zu fördern. Hierfür ist aber die Fähigkeit zur Selbstreflexion relevant, um potenzielle Wissenslücken zu identifizieren, wohingegen ein Verständnis der Epistemologie bei der Auswahl und Bewertung neuer Wissensquellen notwendig ist. Das naturwissenschaftliche Grundwissen ist an dieser Stelle als Grundgerüst vonnöten, um neues Wissen in bestehendes integrieren zu können. Kritisches Denken stellt also keine Abkehr von der fachinhaltlichen Vermittlung dar, sondern vielmehr eine Neuausrichtung der fachlichen Vermittlung mit dem Ziel der maximalen Anwendbarkeit.

Auch für Lehrende spielt Kritisches Denken eine relevante Rolle, besonders im Vermittlungs- und Bewertungskontext. Die Didaktik setzt sich im Kern mit Menschen auseinander, welche verschiedene Voraussetzungen, soziale Einbindungen, Herkünfte, Identitäten, Motivationen und emotionale Zustände mitbringen. Diese Vielfalt in einen pädagogisch sinnvollen Kontext zu setzen, beinhaltet das Potenzial, gezielter auf die Bedürfnisse der*s Einzelnen einzugehen, fachspezifische Inhalte verständlicher aufzubereiten und diese didaktisch zielgerichteter zu transformieren (Reiners, 2022). Kritisches Denken ermöglicht es Lehrkräften, zwischen kognitiven und affektiven Faktoren abzuwägen und so die individuelle Förderung im Chemieunterricht zu verbessern. Dies ist besonders relevant, da emotionale und motivationale Aspekte häufig das Lernverhalten und den Lernfortschritt beeinflussen.

Besondere Bedeutung erlangt Kritisches Denken bei Leistungsbewertungen und in Diagnosesituationen, in denen kognitive Verzerrungen eine Rolle spielen können. Typische Verzerrungen sind unter anderem der *Rosenthal-Effekt*, selbsterfüllende Prophezeiungen durch Erwartungen an die Leistung, der *Kontrast-Effekt*, bei dem Leistungsunterschiede innerhalb einer Gruppe die Beurteilung einzelner Lernender verzerren, oder der *Inferenz-Effekt*, bei dem Lehrkräfte eigene Persönlichkeitsmerkmale auf die Lernenden projizieren oder diese mit gegenteiligen Eigenschaften versehen (Groß, 2022; Kretschmer & Stary, 1998; Sämmer, 1999; Wengert, 2006). Diese Ver-

zerrungen gefährden die Objektivität und Fairness bei der Leistungsbeurteilung, weshalb Kritisches Denken für eine reflektierte und möglichst unvoreingenommene Bewertung unerlässlich ist. Des Weiteren kann Kritisches Denken und insbesondere die Selbstreflexion zur Selbstregulation und so zur Verbesserung des Unterrichts beitragen, indem eigene Urteile und Unterrichtsstrategien hinterfragt oder die Auswahl von Material und Methoden überdacht werden. Die Fähigkeit, über die eigenen Denkmuster nachzudenken, unterstützt die Lehrkräfte dabei, sowohl kognitive Verzerrungen zu vermeiden als auch die emotionalen und motivationalen Bedürfnisse der Lernenden besser zu verstehen. Dies ermöglicht es, empathisch auf Lernende zu reagieren und eine förderliche Lernatmosphäre zu schaffen.

Abschließend ist Kritisches Denken für Lehrkräfte ein entscheidender Faktor zur kontinuierlichen professionellen Weiterentwicklung. Indem der eigene Unterricht selbstständig konstant überprüft wird, wird nicht nur der Unterricht verbessert, sondern auch die Qualität der naturwissenschaftlichen Bildung langfristig gestärkt.

4 Empirischer Teil – Methodik

Ziel dieses Projekts war es, Kritisches Denken nicht nur neu zu definieren und konzeptionell zu fundieren, sondern das Konstrukt auch in seiner Mehrdimensionalität greifbar und operationalisierbar zu machen.

Der ursprüngliche Entwurf dieser Arbeit verfolgte zunächst das Ziel, Ansätze zur expliziten Förderung Kritischen Denkens zu entwickeln. Während der Erstkonzeption zeigte sich jedoch, dass keine geeigneten Instrumente existierten, die der Komplexität und Mehrdimensionalität dieses Konstrukts gerecht werden konnten. Vor diesem Hintergrund wurde die Ausrichtung des Projekts angepasst und die Entwicklung eines Instrumentariums, welches Kritisches Denken sowohl durch Selbst- als auch durch Fremdeinschätzung in seiner Komplexität valide erfassen kann, ins Zentrum des empirischen Erkenntnisinteresses gesetzt.

In diesem Kapitel wird die Grundausrichtung des empirischen Vorgehens dargelegt sowie die Datenerhebungs- und Auswertungsmethoden differenziert beschrieben, um eine bestmögliche Transparenz der einzelnen Schritte zur Wahrung der Güte dieser Arbeit zu gewährleisten.

4.1 Konzeption des empirischen Forschungsprojekts

Wie das vorangegangene Kapitel (vgl. Kap. 3) aufgezeigt hat, stellt Kritisches Denken ein relevantes Bildungsziel dar, welchem insbesondere im Hinblick auf die Naturwissenschaften eine besondere Bedeutung zukommt. Dennoch bleibt neben der Definition des Konstrukts für die Naturwissenschaftsdidaktik die Frage nach der Operationalisierbarkeit offen, um Kritisches Denken auch zielgerichtet adressieren und auf lange Sicht fördern zu können.

Aufgrund der Komplexität des Konstrukts des Kritischen Denkens und des Mangels an hinreichenden Instrumenten, die der Neudefinition gerecht werden, lässt sich der empirische Teil dieser Arbeit als stark explorativ charakterisieren, weswegen schwerpunktmäßig qualitative Instrumente der Datenerhebung Anwendung finden (Lamnek & Krell, 2016). Dabei wird auf Methoden der qualitativen Sozialforschung zurückgegriffen, da diese geeignet sind, den dynamischen und multiperspektivischen Charakter des Kritischen Denkens angemessen zu erfassen und eine verstehend-interpretative Rekonstruktion zu ermöglichen (Döring & Bortz, 2016).

Das explorative Vorgehen erlaubt eine große Flexibilität und eine offene Herangehensweise an den Forschungsprozess, um diesen dynamisch an neue Erkenntnisse anzupassen und neue Perspektiven auf das Thema zu entwickeln (Blumer, 1980). Diese Flexibilität ist jedoch nicht ungerichtet, sondern fokussiert sich im Laufe der Untersuchung, wobei die „Orientierung der Erhebungsinstrumente bzw. deren Auswahl und Anwendung an der Problemstellung und der sozialen Realität" erfolgen (Lamnek & Krell, 2016, S. 37). In Konsequenz bauen die Studien dieses Forschungsprojekts in Anlehnung an das *Prinzip der Zirkularität und Flexibilität* (Döring & Bortz, 2016) iterativ aufeinander auf, wobei die Erkenntnisse einer Studie für die Konzeption oder Überarbeitung der nächsten Studie dienten.

Die handlungsleitende Forschungsfrage für den empirischen Teil dieses Forschungsprojekts wird wie folgt formuliert:

Wie kann Kritisches Denken als Konstrukt im Kontext der Chemiedidaktik operationalisiert werden?

Zur Beantwortung dieser Forschungsfrage und als Orientierung für die Exploration dienen die im Weiteren aufgeführten Untersuchungsfragen:

1. Welche verschiedenen Perspektiven nehmen angehende Chemielehrkräfte bei der Bewertung von gesellschaftlichen Fragestellungen mit naturwissenschaftlichem Bezug (SSI) ein?
2. Welches Verständnis haben angehende Chemielehrkräfte von der Begrifflichkeit des Kritischen Denkens?
3. Inwieweit findet Selbstreflexion bei angehenden Chemielehrkräften bei der Bewertung von gesellschaftlichen Fragestellungen mit naturwissenschaftlichem Bezug statt?

Die Untersuchungsfragen beinhalten je nach Studie und Instrument weitere untergeordnete Untersuchungsfragen, welche der Beantwortung der übergeordneten Untersuchungsfrage dienen und in der Einzelbetrachtung der Studien aufgeführt werden.

Da es das Ziel der empirischen Untersuchung ist, Kritisches Denken operationalisierbar zu machen, müssen in diesem Zuge ebenfalls die Gelingensbedingungen formuliert werden, welche für diese Operationalisierung gelten. Das Instrument muss:

- unabhängig vom Fachwissen sein, da Kritisches Denken als multiperspektivische Denkpraxis und nicht das Fachwissen gemessen werden soll,
- gleichzeitig Zugriff auf mehrere Dimensionen des Denkens ermöglichen, die für eine Anwendung beispielsweise in Bewertungsprozessen relevant sind,
- eine Fremdeinschätzung der Fähigkeit zum Kritischen Denken ermöglichen, welche mit einer Selbsteinschätzung trianguliert werden kann,
- kontextunabhängig sein, um in unterschiedlichen fachdidaktischen Kontexten und Themenbereichen anwendbar zu sein, und
- sowohl die affektive als auch die kognitive Dimension systemisch erfassen.

Da Kritisches Denken als multiperspektivische Denkpraxis verstanden werden kann, die besonders in Entscheidungs- und Bewertungssituationen von hoher Relevanz ist, wurde als Ausgangspunkt der Exploration die Untersuchung von Bewertungsperspektiven bei der Thematisierung von SSI gewählt. Gerade SSI bieten aufgrund ihrer kontroversen Natur, ihrer Uneindeutigkeit, ihrer ethisch-moralischen Anteile sowie ihrer lebensweltlichen Bedeutung sowohl das Potenzial als auch die Notwendigkeit, Kritisches Denken anzuwenden (Bybee, 1997; Gunn et al., 2008; Kolstø et al., 2006; Osborne, 2014; Rafolt et al., 2019; Yacoubian, 2015; Zeidler, 2014).

Ein Kernstück des explorativen Vorgehens stellt dabei die veränderte Ausrichtung der von den Teilnehmenden erwarteten Antworten dar. Angelehnt an Clough (2008), Dewey (1933) sowie Paul und Elder (2014) kann angenommen werden, dass sich die Fähigkeit zu denken in der Fähigkeit widerspiegelt, sich selbst Fragen zu stellen. Dieser Annahme folgend stellt die Formulierung von Fragen innerhalb gegebener Kontexte eine Möglichkeit dar, die Dimensionen des Denkens der Teilnehmenden zu ergründen. Auf diese Weise wird das Potenzial zu denken über den Abruf von (Fach-)Wissen gestellt, da der Formulierung von Fragen eine potenzielle Aneignung von Wissen folgen kann.

Eine weitere Ausrichtung liegt in der Ergründung der Fähigkeit zur Selbstreflexion, da diese im Kontext des Kritischen Denkens als eine wesentliche metakognitive Fähigkeit gilt (vgl. Kap. 3.3). Neben dem Fachwissen und dem epistemologischen Verständnis stellt sie die dritte große Facette Kritischen Denkens in der Naturwissenschaftsdidaktik dar, die jedoch einer vertiefenden Betrachtung bedarf (vgl. Kap. 3.5). Darüber hinaus spielt das Verständnis von Kritischem Denken eine zentrale Rolle,

da es nicht nur für die Anwendung in Bewertungssituationen entscheidend ist, sondern auch für die Fähigkeit angehender Lehrkräfte, das Konzept im schulischen Kontext explizit zu vermitteln. Den zukünftigen Lehrkräften kommt damit im Vermittlungsprozess eine essentielle Rolle als Multiplikator*innen naturwissenschaftlichen Wissens zu (Reiners, 2022), weswegen diese im Zentrum des Erkenntnisinteresses stehen. Nur ein klares Verständnis der Begrifflichkeit ermöglicht eine langfristige, zielgerichtete Vermittlung, die nicht auf individuellen, subjektiven Definitionsansätzen basiert.

Die explorative Ausrichtung dieses Forschungsprojekts erfordert eine flexible und zugleich präzise Methodenauswahl, um den komplexen und multiperspektivischen Charakter des Kritischen Denkens angemessen erfassen zu können. Nachdem die initialen Überlegungen sowie die leitenden Forschungs- und Untersuchungsfragen dargelegt wurden, ist es nun notwendig, die eingesetzten Methoden zur Datenerhebung detailliert zu erläutern, um die Operationalisierung des Kritischen Denkens im Rahmen dieses Projekts nachvollziehbar zu machen.

4.2 Methoden der Datenerhebung

Um die Forschungsfrage zu beantworten und sich explorativ der Operationalisierung von Kritischem Denken zu nähern, kamen unterschiedliche Datenerhebungsinstrumente zum Einsatz. Dies lag in dem Anspruch begründet, die Gütekriterien qualitativer Forschung einzuhalten und vor allem eine hohe Validität zu gewährleisten, da diese nach Petri (2014) das wichtigste Gütekriterium für qualitative und insbesondere fachdidaktische Forschung darstellt. Aufgrund dessen wurden sowohl eigene Werkzeuge erprobt als auch bereits etablierte Testinstrumente eingesetzt, mit dem Ziel, einen Abgleich für die Ausrichtung der eigenen Instrumente vornehmen zu können. Diese wurden vor ihrem Einsatz stets mit Teilnehmenden pilotiert, deren Voraussetzungen mit der Zielgruppe vergleichbar waren, worauf an späterer Stelle detaillierter eingegangen wird.

Insgesamt wurde ein Mixed-Method-Ansatz verfolgt, welcher sowohl qualitative Methoden als auch quantitative Erhebungsformen in sich kombiniert (Döring & Bortz, 2016; Hagenauer & Gläser-Zikuda, 2022). Die quantitativen Methoden liefern dabei strukturierte Daten, welche im Zusammenspiel mit den qualitativen Befunden eine tiefere Interpretation der Ergebnisse ermöglichen. Um beide Perspektiven gleichwertig miteinzubeziehen, wurde ein paralleles Design gewählt, das eine zeitgleiche, aber unabhängige Erhebung und Analyse beider Ansätze vorsieht. Die abschließende

Zusammenführung der Erkenntnisse ermöglicht eine wechselseitige Validierung und trägt zu einem tieferen Verständnis des Untersuchungsgegenstands bei (Abbildung 15) (Creswell, 2015; Hagenauer & Gläser-Zikuda, 2022; Teddlie & Tashakkori, 2009).

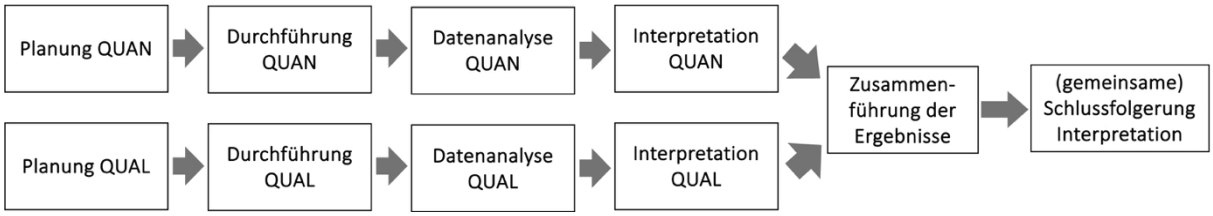


Abbildung 15: Paralleles Mixed-Methods-Design (QUAL + QUAN) (eigene Darstellung basierend auf Hagenauer & Gläser-Zikuda, 2022, S. 256; in Anlehnung an Creswell, 2015, S. 56; Teddlie & Tashakkori, 2009, S. 152)

Die hier durchgeführte Triangulation ist durch die Kombination verschiedener Methodologien geprägt (Denzin, 2017; Lamnek & Krell, 2016) und lässt sich nach Flick (2011, S. 12) als „Einnahme unterschiedlicher Perspektiven auf einen untersuchten Gegenstand" beschreiben. Dabei kam sowohl die sogenannte *Between-Method-Triangulation* zur Kreuzvalidierung und Steigerung der externen Validität zum Einsatz als auch die *Within-Method-Triangulation*, welche primär das Ziel der Reliabilität und internen Konsistenz verfolgt (Lamnek & Krell, 2016).

Konkret strebte die Between-Method-Triangulation (Flick, 2011) bezogen auf das Forschungsinteresse an, ein umfassenderes Bild der Bewertungsstrategien, Einflüsse und des Verständnisses von Kritischem Denken von Lehramtsstudierenden zu gewinnen, aber auch einen Vergleich zwischen Fremd- und Selbsteinschätzung zu ermöglichen. Hierzu wurden sowohl Interviews geführt als auch Fragebögen eingesetzt, wobei die Interviews primär die Fremdeinschätzung ermöglichen sollten und weite Teile der Fragebögen der Selbsteinschätzung dienten. Da die Selbsteinschätzungen durch subjektive Verzerrungen, wie soziale Erwünschtheit oder Selbsttäuschung, beeinflusst sein können, bietet ein Abgleich mit der Fremdeinschätzung eine externe und objektivere Sichtweise (Moosbrugger & Kelava, 2020b). Gleichzeitig erfassen beide Einschätzungen oft unterschiedliche Aspekte eines Verhaltens, wobei die Selbsteinschätzung stärker auf inneren Erfahrungen basiert und die Fremdeinschätzung von außen beurteilt. Die Kombination bietet daher, neben dem Erkennen von Diskrepanzen, einen deutlichen Mehrwert zum Erhalt eines vollständigeren Bildes.

Die Within-Method-Triangulation (Glaser & Strauss, 1965; Lamnek & Krell, 2016) fand durch die Variation von Items, die den gleichen Untersuchungsgegenstand adressierten, und der zusätzlichen Implementation bereits etablierter Werkzeuge in

der quantitativen sowie der Untersuchung verschiedener Vergleichsgruppen bei der qualitativen Datenaufnahme statt. Ziel war die Steigerung der Zuverlässigkeit der Instrumente und damit einhergehend eine höhere Aussagekraft.

Insgesamt wurden im Rahmen dieses Forschungsprojekts fünf Studien, benannt mit den Kürzeln S1 bis S5, an der Universität zu Köln durchgeführt. Die Erhebungen der Studien fanden im Zeitraum des Sommersemesters 2022 bis Wintersemester 2023/24 statt. Studie S1 und S2 bilden dabei die Vorstudien und dienten der Einarbeitung und Erprobung verschiedener Instrumente, wohingegen die Studien S3 bis S5 zur weiteren Untersuchung und Konsolidierung beitrugen (Abbildung 16). Die Teilnehmenden der Studien waren ausnahmslos Studierende des Chemielehramts der Universität zu Köln, wobei die Stichproben verschiedenen chemiedidaktischen Seminaren entstammen.

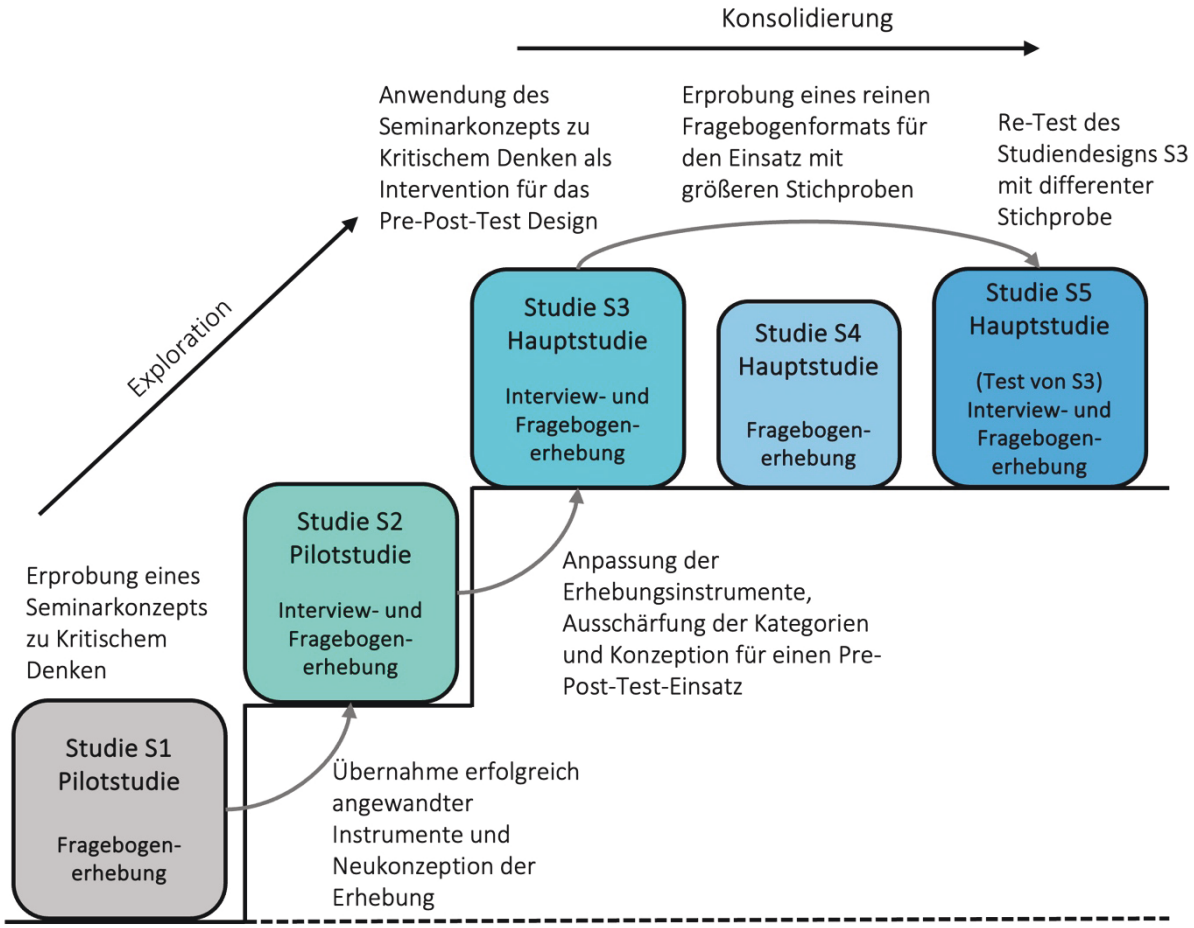


Abbildung 16: Überblick über den Aufbau des Forschungsprojekts. Entwicklung eines Instruments zur Darstellung der Dimensionen des Kritischen Denkens (eigene Darstellung in Anlehnung an Schumacher 2015, S. 53)

Eine Übersicht über die durchgeführten Erhebungen der einzelnen Studien sowie ihre Stichprobengröße kann Tabelle 4 entnommen werden. S1 bis S5 bezeichnen

dabei die jeweiligen Studien, wohingegen die nach dem Punkt folgenden Buchstaben die Art der Erhebung spezifizieren (F für Fragebogen, I für Interview). Die schließenden Zahlen, welche als Suffixe der Spezifizierung folgen, beschreiben die Nummer der Erhebung, wobei 0 für die Pilotierung und die positiven ganzen Zahlen für die Haupterhebungen stehen.

Tabelle 4: Übersicht über die einzelnen Studien des Forschungsprojekts und die einzelnen Stichproben

Studien		Fragebogenerhebungen			Interviewerhebungen		
Vorstudien	Studie 1	S1.F1	S1.F2	S1.F3			
		15	12	15			
	Studie 2	S2.F1			S2.I0	S2.I1	
		10			3	10	
Hauptstudien	Studie 3	S3.F0	S3.F1	S3.F2	S3.I1	S3.I2	S3.I3
		4	16	16	6	10	3
	Studie 4	S4.F1	S4.F2				
		17	14				
	Studie 5	S5.F1	S5.F2		S5.I1	S5.I2	S5.I3
		12	12		7	7	5
		Gesamt:		**143**	**Gesamt:**		**51**

Fragebögen
Interviews

Insgesamt wurden im Rahmen des Forschungsprojekts 143 vollständige Fragebögen und 51 Interviews erhoben und anschließend ausgewertet.

4.2.1 Datenschutz

Die Teilnehmenden am Forschungsprojekt wurden für die Gewährleistung der informierten Einwilligung und freiwilligen Teilnahme vor der Datenaufnahme und -verarbeitung schriftlich und mündlich aufgeklärt. Diese Aufklärung erfolgte über

- den Zweck und Aufbau des Forschungsprojektes,
- die zu erwartende Dauer der Teilnahme und das Vorgehen,
- das Recht auf eine Ablehnung der Teilnahme und frühzeitige Beendigung der Teilnahme ohne Konsequenzen,
- den voraussichtlichen Erkenntnisgewinn durch die Forschungsarbeit,
- die Gewährleistung von Anonymität/Pseudonymität, Vertraulichkeit und deren Grenzen sowie
- die Möglichkeiten der Kontaktaufnahme bei Fragen zum Forschungsvorhaben oder den Rechten als Forschungsteilnehmende (Döring & Bortz, 2016).

Innerhalb der Fragebogenerhebung wurde vor dem Ausfüllen durch die Teilnehmenden ein individueller sechsstelliger Code generiert, welcher statt des Namens verwendet wurde. Dieser basierte auf drei Fragen, um eine Rekonstruktion in späteren Fragebögen zu ermöglichen. Solange lediglich der Fragebogen ausgefüllt wurde, konnte so zu einem hohen Grad die Anonymität gewährleistet werden, was einer Verfälschung des Antwortverhaltens durch beispielsweise soziale Erwünschtheit entgegenwirken kann (Busker, 2014; Döring & Bortz, 2016). Die Pseudonymisierung trug in diesem Kontext ebenfalls dazu bei, eine möglichst unvoreingenommene Auswertung zu gewährleisten, da die Datenerhebung größtenteils in Seminaren stattfand, in denen die untersuchende Person zugleich als Lehrkraft tätig war, wodurch ein persönlicher Bezug zwischen Teilnehmenden und Lehrkraft bestand. Die genauen Rahmenbedingungen können den Studienbeschreibungen in Kapitel 4.2.2 und 5 entnommen werden. Der Code wurde überdies zu Triangulationszwecken genutzt, um bei der Teilnahme an der Interviewstudie eine Verknüpfung beider Datensätze zu ermöglichen (Moosbrugger & Brandt, 2020c). Alle im Rahmen des Forschungsprojektes erhobenen Daten, welche potenziell zur Identifikation der Teilnehmenden führen könnten, wurden im Nachgang geschwärzt bzw. pseudonymisiert. Zusätzlich wurden alle von den Teilnehmenden generierten Codes mithilfe eines Chiffrierungsschlüssel abermals verschlüsselt, um eine Rückverfolgung über die Erstellungsfragen zu erschweren. Im Rahmen von Vorträgen und Publikationen wurden ausschließlich anonymisierte und chiffrierte Datensätze verwendet.

Die Datenaufnahme in den Interviewerhebungen fand ausschließlich durch Aufnahmegeräte mit lokalem Datenspeicher statt. Bei der darauffolgenden Vortranskription wurde das KI-Tool *f4x* verwendet, welches jedoch DSGVO-konform auf verschlüsselten deutschen Servern operiert. Die hochgeladenen Daten werden dabei nach der Verarbeitung und erfolgtem Download gelöscht und bleiben bis dahin maximal 28 Tage abrufbar. Im Zuge der Nachtranskription wurden personenbezogene Daten anonymisiert.

Die Datenaufnahme der Fragebögen erfolgte über die Online-Umfrage-Applikation *LimeSurvey* (Limesurvey GmbH, o. J.), welche seitens der Universität zu Köln bereitgestellt wurde und innerhalb des Verantwortungsbereiches des Regionalen Rechenzentrums der Universität zu Köln (RRZK) liegt. Die mittels LimeSurvey erhobenen Daten wurden vor der Speicherung ebenfalls auf Anonymität überprüft und nötigenfalls nachträglich anonymisiert. Alle über die Plattform LimeSurvey aufgenommenen Daten werden nach Beendigung der Befragung innerhalb von 24 Monaten vom Server gelöscht.

Alle Forschungsdaten, welche im Rahmen des Projektes erhoben wurden, werden auf Grundlage von § 13 (Archivierung) der Leitlinien zur guten wissenschaftlichen Praxis der Universität zu Köln (Universität zu Köln, 2022) zehn Jahre anonymisiert aufbewahrt und anschließend gelöscht.

4.2.2 Stichprobenbeschreibung der Erhebungen

Alle Erhebungen des Forschungsprojekts wurden in verschiedenen chemiedidaktischen Seminaren der Universität zu Köln durchgeführt. Die Seminare stellten häufig Kombinationsveranstaltungen der Studienprofile

- Lehramt an Gymnasien und Gesamtschulen,
- Lehramt an Haupt-, Real-, Sekundar- und Gesamtschulen sowie
- Lehramt für sonderpädagogische Förderung

dar, weswegen schwerpunktmäßig heterogene Stichproben bezüglich der Studienprofile auftraten. Ein tabellarischer Überblick über die einzelnen Stichproben der Erhebungen des Forschungsprojekts ist Tabelle 5 am Ende des Kapitels zu finden.

Bevor es an dieser Stelle zur weiteren detaillierten Beschreibung der Stichproben kommt, soll zur Vermeidung von Missverständnissen herausgestellt werden, dass die Bezeichnung der Gesamtstichprobe basierend auf den aktuellen Richtlinien der APA (American Psychological Association, 2020) mit dem Großbuchstaben *N* erfolgt, welcher die Gesamtanzahl Teilnehmenden einer Erhebung repräsentiert. Teilstichproben werden demzufolge in dieser Arbeit mit dem Kleinbuchstaben *n* beschrieben. Die in der quantitativen Statistik gebräuchliche Verwendung des Großbuchstabens *N* für die Gesamtpopulation und *n* für die gezogene Stichprobe findet hier keine Anwendung, da hier im Rahmen des qualitativen Forschungsparadigmas der aktuellen Studien ausschließlich auf die Stichprobengröße der Teilnehmenden rekurriert wird und keine direkten Rückschlüsse auf eine übergeordnete Population beabsichtigt sind.

Die initialen Erhebungen des Forschungsprojektes fanden im Rahmen des *Seminars zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP)* im Sommersemester 2022 statt, welches der Untersuchende als Lehrperson gestaltete. Bei diesem Seminar handelte es sich um ein obligatorisches Seminar im Bachelorstudiengang Lehramt Chemie für die verschiedenen Schulformen, welches im Aufbaumodul *Ausgewählte Aspekte des Chemieunterrichts* der jeweiligen Modulhandbücher verortet ist

(Institut für Chemiedidaktik (Hrsg.), 2023b, 2023a; Math.-Nat. Fakultät der Universität zu Köln (Hrsg.), 2018). Die Stichproben aller Erhebungen von Studie 1 entstammten derselben Gruppe von Seminarteilnehmenden, wobei die Stichprobengröße leicht variierte, da die Teilnehmendenzahl nicht in jeder Sitzung, in der erhoben wurde, konstant war. Die Stichprobengröße belief sich auf $N = 15$ in der Fragebogenerhebung S1.F1, auf $N = 12$ in der Fragebogenerhebung S1.F2 sowie auf $N = 15$ in der abschließenden Fragebogenerhebung S1.F3. Ab der dritten Fragebogenstudie wurden auch weitere personenbezogene Daten, wie das Studienprofil in Bezug auf die Schulform, die Anzahl an Fachsemestern und die Zweitfächer erhoben. Die am stärksten vertretene Gruppe bildeten 13 Teilnehmende mit dem Studienprofil Lehramt für Gymnasien und Gesamtschulen, wohingegen Lehramtsstudierende für Haupt-, Real- und Gesamtschule sowie Sonderpädagogik jeweils nur einmal vorkamen.

Im Wintersemester 2022/2023 wurde erstmals im Rahmen der Studie S2 ein Interviewleitfaden für die geplanten Interviewerhebungen pilotiert (S2.I0). Die Pilotierungsstichprobe bestand aus $N = 3$ Studierenden, wobei ein*e Teilnehmende des Lehramts das Studienprofil Sonderpädagogik und zwei Teilnehmende für Haupt-, Real-, Sekundar- und Gesamtschulen studierten. Die Teilnehmenden stammten aus der Chemie-Vorlesung des Moduls *Mathematisch-Naturwissenschaftliche Grundlegung (MNF-B)*, welches von Lehramtsstudierenden anderer Naturwissenschaften als Ergänzungsmodul besucht wird (Institut für Chemiedidaktik (Hrsg.), 2023a, 2023b; Math.-Nat. Fakultät der Universität zu Köln (Hrsg.), 2018). Damit konnte sichergestellt werden, dass die Teilnehmenden nicht im späteren Verlauf Teil der eigentlichen Stichprobe von Chemielehramtsstudierenden wurden, aber gleichzeitig eine für die Pilotierung vergleichbare fachliche Eignung seitens der Teilnehmenden vorlag. Zwei Teilnehmende befanden sich im ersten Semester, eine Person im vierten Semester. Die auf die Pilotierung folgende Interview- und Fragebogenerhebung (S2.F1 & S2.I1) fand mit $N = 10$ Studierenden des Chemielehramts statt. Diese setzte sich aus sechs Studierenden im Bachelorstudiengang und vier Studierenden im Masterstudiengang Lehramt Chemie zusammen. Die Bachelorstudierenden entstammten dem Seminar zum Modul *Grundlegende Aspekte der Fachdidaktik (B2 HR/SP B3 GG)* (Institut für Chemiedidaktik (Hrsg.), 2023b, 2023a; Math.-Nat. Fakultät der Universität zu Köln (Hrsg.), 2018), wobei alle das Studienprofil für Gymnasien und Gesamtschulen aufwiesen. Die Studierenden hier waren durchschnittlich im dritten Fachsemester, ein*e Teilnehmende*r hatte ein naturwissenschaftliches Fach als Zweitfach. Die Masterstudierenden entstammten hingegen dem Seminar zu *Ausgewählten Themengebieten des Chemieunterrichts (M1 HR/GG)* (Institut für Chemiedidaktik (Hrsg.), 2023c;

Math.-Nat. Fakultät der Universität zu Köln (Hrsg.), 2021) und waren ebenfalls im Studienprofil für Gymnasien und Gesamtschulen verortet. Die Studierenden waren durchschnittlich im dritten Fachsemester des Masters und zwei von ihnen studierten ein naturwissenschaftliches Zweitfach. Beide Seminare, aus denen die Teilnehmenden stammten, wurden von anderen Dozierenden und nicht vom Untersuchenden selbst gestaltet.

Die Erhebungen der dritten Studie S3 fanden im Sommersemester 2023 statt. Hier wurde ein neuer, kombinierter Fragebogen eingesetzt, welcher mit $N = 4$ Masterstudierenden aus dem Seminar *Vertiefende Aspekte der Chemiedidaktik (M1 HR/SP/GG)* (Institut für Chemiedidaktik (Hrsg.), 2023c, 2023d; Math.-Nat. Fakultät der Universität zu Köln (Hrsg.), 2021) pilotiert wurde (S3.F0). Alle Teilnehmenden waren im Studienprofil Lehramt Chemie für Gymnasien und Gesamtschulen verortet und waren fast alle im ersten Fachsemester des Masters. Hier wies keiner der Befragten ein naturwissenschaftliches Zweitfach auf. Die Haupterhebungen S3.F1&2 sowie S3.I1&2, welche beide im Pre-Post-Test-Design gehalten waren, wurden mit Bachelorstudierenden des Chemielehramts im *Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP)* durchgeführt, welches vom Untersuchenden selbst als Lehrperson betreut wurde. An der Fragebogenerhebung nahmen $N = 16$ Studierende teil, zwölf davon mit Studienprofil für Gymnasien und Gesamtschulen, zwei für sonderpädagogische Förderung und eine Person für Haupt-, Real-, Sekundar- und Gesamtschulen. Die Studierenden waren durchschnittlich im sechsten Fachsemester und fünf von ihnen wiesen ein naturwissenschaftliches Zweitfach auf. Von diesen 16 Studierenden nahmen zehn an der Interviewstudie S3.I1&2 teil. Vier davon hatten bereits an der Interviewstudie S2.I1 teilgenommen, weshalb mit diesen nur ein abschließendes Post-Interview in S3.I2 geführt wurde. Dies war dadurch möglich, dass sich der Interviewleitfaden nach seiner Pilotierung S2.I0 nicht mehr grundlegend verändert hatte und der zeitliche Abstand zur Erhebung S2.I1 nicht zu groß war. Damit war eine Vergleichbarkeit zu den sechs Studierenden gegeben, welche an der Pre-Erhebung S3.I1 während des Seminars teilnahmen, weshalb auf diese Weise eine Pre-Post-Stichprobe von $N = 10$ zustande kam. Die Stichprobe setzte sich aus acht Studierenden mit Studienprofil für Gymnasien und Gesamtschulen sowie zwei mit Profil für sonderpädagogische Förderung zusammen. Die Teilnehmenden an der Interviewerhebung waren durchschnittlich im fünften Semester, zwei Studierende studierten ein naturwissenschaftliches Zweitfach.

Die vierte Studie S4 war eine reine Fragebogenstudie und fand im Rahmen des chemiedidaktischen Seminars *Einführung in das wissenschaftliche Arbeiten* im Basismodul

Grundlegende Aspekte der Fachdidaktik (B2 HR/SP B3 GG) (Institut für Chemiedidaktik (Hrsg.), 2023b, 2023a; Math.-Nat. Fakultät der Universität zu Köln (Hrsg.), 2018) statt, welches ebenfalls vom Untersuchenden selbst konzipiert und geleitet wurde. Dieses wurde im Sommersemester 2023 als Blockveranstaltung für zwei Gruppen gehalten. Die beiden Gruppen unterschieden sich hinsichtlich der im Fragebogen verwendeten Kontexte, waren jedoch ansonsten identisch. Die Stichprobe der ersten Gruppe S4.F1 bestand aus $N = 17$ Teilnehmenden, wovon 13 das Studienprofil für Gymnasien und Gesamtschulen, drei für Haupt-, Real-, Sekundar- und Gesamtschulen und zwei für sonderpädagogische Förderung vorwiesen. Sechs Studierende hatten ein naturwissenschaftliches Zweitfach und befanden sich durchschnittlich im dritten Fachsemester. Die zweite Gruppe S4.F2 bestand aus $N = 14$ Teilnehmenden, zehn davon mit Studienprofil für Gymnasien und Gesamtschulen, zwei für sonderpädagogische Förderung und ein*e Befragte*r für Haupt-, Real-, Sekundar- und Gesamtschulen. Auch hier waren die Studierenden durchschnittlich im dritten Fachsemester, wohingegen hier sieben Teilnehmende ein naturwissenschaftliches Zweitfach hatten.

Die abschließenden Erhebungen der Studie S5 fanden im Wintersemester 2023/2024 mit Studierenden des Masterstudiengangs Lehramt Chemie im Seminar *Vertiefende Aspekte der Chemiedidaktik (M1 HR/SP/GG)* statt. Der Untersuchende trat hier als Gastdozent in einem ansonsten von einem anderen Dozierenden der Chemiedidaktik konzipierten Seminar in Erscheinung. Hier nahmen $N = 12$ Studierende an der Pre-Post-Fragebogenerhebung S5.F1&2 teil, davon neun Studierende mit Studienprofil für Gymnasien und Gesamtschulen, zwei für sonderpädagogische Förderung und ein*e Studierende*r für Haupt-, Real-, Sekundar- und Gesamtschulen. Die Studierenden waren durchschnittlich im ersten Fachsemester und sechs von ihnen wiesen ein naturwissenschaftliches Zweitfach auf. An der abschließenden Pre-Post-Interviewerhebung nahmen sieben Studierende teil, welche ebenfalls durchschnittlich im ersten Fachsemester des Masters waren. Vier Teilnehmende studierten für das Lehramt an Gymnasien und Gesamtschulen, eine Person für Haupt-, Real-, Sekundar- und Gesamtschulen und zwei für sonderpädagogische Förderung. Hier wiesen zwei Studierende ein naturwissenschaftliches Zweitfach auf.

Eine umfassendere Übersicht über die Verortung der Bachelormodule der verschiedenen Studienprofile kann den jeweiligen Modulhandbüchern der Universität zu Köln entnommen werden (Institut für Chemiedidaktik (Hrsg.), 2023a, 2023b; Math.-Nat. Fakultät der Universität zu Köln (Hrsg.), 2018). Eine Einordnung der Stellung der Seminare im Studienverlauf des Masterstudiums ist ebenfalls in den zugehörigen

Modulhandbüchern zu finden (Institut für Chemiedidaktik (Hrsg.), 2023c, 2023d; Math.-Nat. Fakultät der Universität zu Köln (Hrsg.), 2021).

Tabelle 5: Tabellarische Übersicht über die einzelnen Stichproben der Forschungsprojektes und ihrer Charakteristika

Erhebung	Modul	Semester	Stichprobengröße (*N*)	Verteilung auf Schulformen	Teilnehmende mit nawi. Zweitfach	⌀ Durchschnittl. Fachsemester
S1.F1	B5 HR/GG B3 SP	SoSe 22	15	Nicht erhoben	Nicht erhoben	Nicht erhoben
S1.F2	B5 HR/GG B3 SP	SoSe 22	12	Nicht erhoben	Nicht erhoben	Nicht erhoben
S1.F3	B5 HR/GG B3 SP	SoSe 22	15	Gym/Ges: 13 HRSG: 1 SoPäd: 1	5	4,3
S2.F1	B2 HR/SP B3 GG & M1 HR/GG	WiSe22/23	B.A.: 6 M.Ed.: 4	Gym/Ges B.A.: 6 Gym/Ges M. Ed.: 4	B.A.: 1 M.Ed.: 2	B.A.: 3 M.Ed.: 3,3
S2.I0	GG-MNF-B	WiSe22/23	3	HRSG: 2 SoPäd: 1	Nicht erhoben	2,0
S2.I1	B2 HR/SP B3 GG & M1 HR/GG	WiSe22/23	B.A.: 6 M.Ed.: 4	Gym/Ges B.A.: 6 Gym/Ges M. Ed.: 4	B.A.: 1 M.Ed.: 2	B.A.: 3 M.Ed.: 3,3
S3.F0	M1 HR/SP/GG	SoSe23	4	Gym/Ges: 4	0	1,5
S3.F1&2	B5 HR/GG B3 SP	SoSe23	16	Gym/Ges: 12 HRSG: 1 SoPäd: 2	5	6,0
S3.I1	B5 HR/GG B3 SP	SoSe23	6	Gym/Ges: 4 SoPäd: 2	1	5,8
S3.I2	B5 HR/GG B3 SP	SoSe23	10 (+ 3 Follow-Up)	Gym/Ges: 8 SoPäd: 2	2	5,1
S4.F1	B2 HR/SP B3 GG	SoSe23	17	Gym/Ges: 13 HRSG: 3 SoPäd: 2	6	3,1
S4.F2	B2 HR/SP B3 GG	SoSe23	14	Gym/Ges: 10 HRSG: 1 SoPäd: 3	7	3,1
S5.F1&2	M1 HR/SP/GG	WiSe23/24	12	Gym/Ges: 9 HRSG: 1 SoPäd: 2	6	1,3
S5.I1&2	M1 HR/SP/GG	WiSe23/24	7 (+ 5 Follow-Up)	Gym/Ges: 4 HRSG: 1 SoPäd: 2	2	1,3

4.2.3 Fragebogenerhebungen

Da sich die Items und Gestaltungen der Fragebögen in Abhängigkeit von den zugehörigen Untersuchungsfragen und aufgrund des iterativen Ansatzes voneinander unterscheiden, wird im Folgenden auf die grundlegende Konzeption der Fragebogenerhebungen eingegangen. Eine detailliertere Betrachtung der einzelnen eingesetzten Items erfolgt in den jeweiligen Kapiteln der Studien.

Die Fragebogenerhebungen dienten vordergründig der Untersuchung der individuellen Sichtweisen und der Selbsteinschätzung der Teilnehmenden (Moosbrugger & Brandt, 2020c; Tiemann & Körbs, 2014). Dabei bot der Einsatz von Fragebögen gleichzeitig die Möglichkeit, eine größere Stichprobe von Teilnehmenden zeitökonomisch zu untersuchen. Zur Vereinfachung der Datenaufnahme wurde die Online-Umfrage-Applikation *LimeSurvey* verwendet, welche eine digitale Bearbeitung der Fragen seitens der Teilnehmenden ermöglichte und die Datenverarbeitung bei der späteren Auswertung erleichterte (Hussy et al., 2010; Limesurvey GmbH, o. J.). Zusätzlich konnte durch die Vostrukturierung des Fragebogens sichergestellt werden, dass bei der Beantwortung der Fragen der vorangegangenen Datenschutzerklärung zugestimmt wurde, welcher eine hohe Relevanz in der Aufnahme und Verarbeitung personenbezogener Daten zukommt (Lamnek & Krell, 2016; Moosbrugger & Brandt, 2020c). Dem Fragebogen ging eine Seite mit der Aufklärung über die Ziele des Fragebogens voran, mit der Bitte, jede Seite des Fragebogens auszufüllen, da nur vollständige Datensätze bei der Auswertung berücksichtigt werden können. Auf diese Weise konnte von einer Freiwilligkeit der Teilnahme und einer informierten Einwilligung bei der weiteren Beantwortung der Fragen ausgegangen werden (Döring & Bortz, 2016). Zum Schutz der Teilnehmenden begann jeder Fragebogen mit einer Anleitung zur Erstellung eines individuellen sechsstelligen Codes, welcher an die Stelle des Namens trat. Diese Pseudonymisierung diente zum einen der Zusicherung der Anonymität beim Ausfüllen des Fragebogens und damit Reduktion der sozialen Erwünschtheit (Busker, 2014), ermöglichte aber zugleich die freiwillige Meldung durch die Teilnehmenden bei eventuell auftretenden Rückfragen (Riese & Reinhold, 2014). Zum anderen diente diese Pseudonymisierung zur Wahrung der Unvoreingenommenheit bei der Auswertung, da die Teilnehmenden auch Studierende im Seminar waren und daher ein persönlicher Bezug zu ihnen vorhanden war. Die personenbezogenen Daten, wie der Lehramtsstudiengang, weitere Fächer neben Chemie, der aktuelle Abschnitt des Studiums und das Semester, wurden unter Berücksichtigung der Fragebogenstruktur am Anfang des Fragebogens nach der Erstellung des Codes erhoben (Tiemann & Körbs, 2014).

Die eigentlichen Items umfassten verschiedene Fragentypen. Die offenen/nicht standardisierten Fragestellungen dienten zur vertieften Einsicht in die Bewertungsstrategien und Einflüsse der Teilnehmenden sowie zur Einsicht in das subjektive Verständnis von Begrifflichkeiten wie *Kritischem Denken*, *Selbstreflexion* oder *kognitiven Verzerrungen*. Dieses Format ermöglicht, dass die Teilnehmenden gemäß ihrer eigenen Vorstellungen und in ihrer Sprachlichkeit antworten können (Porst, 2014; Reinders, 2022a). Damit einhergehend bedarf es aber zugleich einer erweiterten Auswertung der gegebenen Antworten (Riese & Reinhold, 2014). Die geschlossenen/standardisierten Items beinhalteten hauptsächlich diskret gestufte Ratingskalen, welche auch als *Likert-Skalen* bezeichnet werden (Moosbrugger & Brandt, 2020a) und zu den am weitesten verbreiteten Skalen in Fragebögen der empirischen Sozialforschung zählen (Porst, 2014). Die bei den selbsterstellten geschlossenen Items verwendeten Skalen bestanden aus Ratingskalen mit fünf Abstufungen, da fünf bis sieben Abstufungen bei Ratingskalen optimale Eigenschaften hinsichtlich der Validität und Reliabilität aufweisen (Moosbrugger & Brandt, 2020a; Weijters et al., 2010). Zur Konkretisierung sowie besseren Interpretation durch die Teilnehmenden und unter Berücksichtigung der Polarität wurden numerische und verbale Skalenpunktbeschreibungen miteinander kombiniert (Moosbrugger & Brandt, 2020a), wobei die verbalen Beschreibungen stets an das Item angepasst wurden. In diesem Kontext wurden alle Skalenpunkte einer Skala benannt, da dies einen Einfluss auf das subjektive Empfinden und die Zufriedenheit der Teilnehmenden nehmen kann (Johnson et al., 2005; Moosbrugger & Brandt, 2020a; Weng, 2004) und sich die Tendenz zum extremen Urteil verringert (Moors et al., 2014). Die Tendenz zur Mitte kann bei einer fünfstufigen Ratingskala nicht ausgeschlossen werden (Moosbrugger & Brandt, 2020a; Van Herk et al., 2004), jedoch stellt die mittlere Kategorie bei einer unipolaren Skala keine neutrale Antwort dar und die Vorteile einer ungeraden Anzahl an Skalenpunkten überwiegen (Weijters et al., 2010). Eine mindestens fünfstufige Ratingskala erlaubt zudem, die erhobenen Daten als quasi-metrisch zu betrachten, wenn die Abstände zwischen den Antworten als semantisch gleich und durch numerische Wertzuweisung als gleich groß interpretiert werden können (Urban & Mayerl, 2011; Völkl & Korb, 2018). Dadurch entsteht die Möglichkeit, wie bei einer Intervallskala erweiterte statistische Operationen bei der Datenanalyse einsetzen zu können, wie die Berechnung von Streumaßen (ebd.). Die auf dieser Grundlage berechneten Standardabweichungen und Varianzen der eingesetzten Likert-Skalen wurden auf Grundlage von Tabelle 6 eingeschätzt.

Tabelle 6: Einschätzung der Varianz von Rating-Skalen (Iacobucci & Churchill, 2010)

Skalenpunkte	Typischer Varianzbereich
4	0,7 – 1,3
5	1,2 – 2,0
6	2,0 – 3,0
7	2,5 – 4,0
10	3,0 - 7,0

Neben den selbsterstellten Items kamen auch zwei etablierte Werkzeuge aus der Psychologie zum Einsatz. Zum einen fand die kurze Version der *Self-Reflection and Insight Scale* (SRIS) Einzug in den Fragebogen (Silvia, 2021). Dabei handelt es sich um ein psychometrisches Instrument, das entwickelt wurde, um die Selbstreflexion (Self-Reflection) und die Einsicht (Insight), zwei zentrale Aspekte des persönlichen Selbstbewusstseins, zu messen (Grant et al., 2002; Silvia, 2021). Die kurze Version nach Silvia (2021) besteht aus zwei Hauptskalen mit je sechs Items zur Selbstreflexion (SRIS-SR) und Einsicht (SRIS-IN) (Tabelle 7). Dabei handelt es sich um siebenstufige unipolare Ratingskalen, wobei die Skala der Selbstreflexion die Tendenz einer Person misst, regelmäßig über ihre Gedanken, Gefühle und Handlungen nachzudenken und diese zu bewerten, was im Kontext dieses Forschungsprojektes einen deutlichen Mehrwert darstellt. Die Skala der Einsicht bezieht sich auf das Maß an Klarheit und Verständnis, das eine Person über ihre inneren Zustände hat und inwieweit sie ihre Gedanken, Gefühle und Handlungen verstehen kann.

Tabelle 7: Items der SRIS nach Silvia (2021), basierend auf dem SRIS von Grant et al. (2002), die im späteren Verlauf rekodierten inversen Items sind durch (*) gekennzeichnet

Self-Reflection (SRIS-SR)	
SR3	Ich betrachte häufig meine Gefühlswelt.
SR5	Ich nehme mir häufig Zeit, um über meine Gedanken zu reflektieren.
SR6	Ich denke oft über die Art und Weise nach, wie ich mich in bestimmten Situationen fühle.
SR8	Es ist mir wichtig, die Dinge zu bewerten, die ich tue.
SR9	Ich bin sehr daran interessiert, nachzuvollziehen, wie ich denke.
SR10	Es ist mir sehr wichtig, zu verstehen, was meine Gefühle bedeuten.
Insight (SRIS-IN)	
IN2	Ich bin oft verwundert über die Art und Weise, wie ich wirklich über Dinge denke. (*)
IN4	Ich bin mir meiner Gefühle bewusst, kann diese aber oft nicht klar benennen. (*)
IN5	Mein Verhalten erstaunt mich oft. (*)
IN6	Das Nachdenken über meine Gedanken bringt mich durcheinander. (*)
IN7	Ich finde es oft schwierig, meine Gefühlswelt zu verstehen. (*)
IN8	Ich weiß normalerweise, warum ich so fühle, wie ich fühle.

Beide Skalen sind dabei essenziell für ein Verständnis der Selbstregulierung, da sie versuchen, die Fähigkeiten zu erfassen, die für eine bewusste Reflexion, Bewertung und Anpassung relevant sind. Die kurze Version der SRIS stellt dabei die überarbeitete Version der SRIS von Grant et al. (2002) dar, welche jeweils zehn Items pro Skala enthielt, um eine effiziente Implementation in Online-Fragebögen zu ermöglichen (Silvia, 2021). Die ursprüngliche Version der SRIS wurde mit $N = 260$ Teilnehmenden in der ersten Validierungsstudie erprobt und die Test-Retest-Reliabilität mit weiteren $N = 28$ Teilnehmenden überprüft (Grant et al., 2002). Zur weiteren Optimierung eliminierte die kurze Version Items, welche die psychometrische Qualität beeinträchtigten sowie niedrige Diskriminationswerte aufwiesen, und wurde mit $N = 1192$ Teilnehmenden getestet (Silvia, 2021). Sowohl die SRIS nach Grant et al. (2002) als auch die Kurzversion von Silvia (2021) gelten als validierte und etablierte Werkzeuge, welche auch in anderen Studien Verwendung finden (Harrington & Loffredo, 2010; Silvia, 2021; Stein & Grant, 2014). Aufgrund der effizienten Implementationsmöglichkeiten der kurzen SRIS bei gleichbleibender Robustheit des Tests sowie der Passung der Messmethode auf den hier im Forschungsprojekt adressierten Untersuchungsgegenstand, wurden die Items der SRIS im Laufe der Studien in den Fragebogen zur Erhebung integriert und eingesetzt. Da zu diesem Instrument keine Normwerttabellen vorlagen, wurde die Auswertung schwerpunktmäßig deskriptiv vorgenommen, wobei Summenscores berechnet wurden, um die Werte der beiden Skalen miteinander vergleichen zu können.

Ein weiteres und bereits etabliertes Instrument stellt der *Cognitive Reflection Test* (CRT) von Frederick (2005) dar, welches die Fähigkeit einer Person misst, intuitive Antworten zu unterdrücken und stattdessen eine reflektierte, logische Antwort zu geben. Beim Treffen von Entscheidungen kann es zu Heuristiken auf Basis des angesprochenen Systems kommen (vgl. Kap. 2.3.2). Das intuitive System 1 trifft in diesen Fällen eine schnelle Vorentscheidung, welche im Weiteren jedoch nicht vom rationalen System 2 korrigiert wird (Frederick, 2002; Kahneman, 2012). Mit dem Ziel, ein Maß für die kognitive Reflexion zu entwickeln, erprobte Frederick (2005) ein Set bestehend aus drei Fragen, welche eine auf den ersten Blick intuitive, aber falsche Antwort nahelegen und ein bewusstes Nachdenken erfordern (Tabelle 8).

Tabelle 8: Fragen des CRT nach Frederick (2005)

Frage 1	Ein Schläger und ein Ball kosten zusammen 1,10 €. Der Schläger kostet 1,00 € mehr als der Ball. Wie viel kostet der Ball? [Euro:][8]
Frage 2	Wenn 5 Maschinen 5 Minuten benötigen, um 5 Geräte herzustellen, wie lange dauert es, mit 100 Maschinen 100 Geräte herzustellen? [Minuten:]
Frage 3	In einem See ist ein Feld mit Seerosen. Jeden Tag verdoppelt sich die Größe des Feldes. Wenn die Seerosen 48 Tage brauchen, um den gesamten See zu bedecken, wie lange dauert es, bis sie die Hälfte des Sees bedecken? [Tage:]

Damit misst der Test die Erstimpulskontrolle in Entscheidungsprozessen und rekurriert auf die Dual-Process Theory und die Stärke des schnellen und intuitiv agierenden Systems 1 gegenüber dem langsamen, rationalen Systems 2. Teilnehmende mit einem höheren Score scheinen in Entscheidungsprozessen geduldiger zu sein und überlegtere Entscheidungen zu treffen, was nahelegt, dass diese ihre ersten intuitiven Impulse besser kontrollieren und ihre rationale Ebene zielgerichteter ansteuern können. Das Fragenset wurde mit 3428 Teilnehmenden in 35 separaten Studien über einen Zeitraum von 26 Monaten getestet, unter anderem mit Studierenden des MIT und der Princeton University (ebd.). Trotz seiner Einfachheit weist der CRT große Korrelationen mit komplexeren Tests, wie dem Scholastic Achievement Test (SAT) oder dem American College Test (ACT), auf, wobei sich der CRT als einer der besten Prädiktoren für erfolgreiche Entscheidungsprozesse zeigte (ebd.). Aufgrund der einfachen Implementierung und der gleichzeitigen hohen Aussagekraft wurde der CRT im Fragebogen abwechselnd zu den anderen Fragen eingesetzt. Diese Integration fand in Anlehnung an die Erprobungsstudien des CRT statt, in welchen die drei Fragestellungen ebenfalls als Teil eines größeren Tests eingebunden wurden (ebd.). Zur Vermeidung von Anker- und Konsistenzeffekten (Chapman & Johnson, 2002; Moosbrugger & Brandt, 2020b; Salancik & Pfeffer, 1977) wurde die Reihenfolge der drei Fragen in LimeSurvey randomisiert.

Die Auswertung des CRT erfolgte über die Berechnung von Mittelwerten und die Analyse der prozentualen Verteilung der Antworten. Zur Einordnung der empirischen Daten wurden die von Frederick (2005) bereits erhobenen Referenzwerte herangezogen (Tabelle 9).

[8] Die Antworten auf den Cognitive Reflection Test (CRT) von Frederick (2005) sollen an dieser Stelle natürlich nicht vorenthalten werden:
Frage 1: Korrekte Antwort: 0,05 € Intuitive Antwort: 0,10 €
Frage 2: Korrekte Antwort: 5 Minuten Intuitive Antwort: 100 Minuten
Frage 3: Korrekte Antwort: 47 Tage Intuitive Antwort: 24 Tage

Tabelle 9: Vergleichswerte für die Auswertung des CRT (Frederick, 2005, S. 29)

		Percentage scoring 0, 1, 2 or 3				
Locations at which data were collected	*Mean CRT score*	*"Low" 0*	*1*	*2*	*"High" 3*	*N =*
Massachusetts Institute of Technology	2.18	7%	16%	30%	48%	61
Princeton University	1.63	18%	27%	28%	26%	121
Boston fireworks display[a]	1.53	24%	24%	26%	26%	195
Carnegie Mellon University	1.51	25%	25%	25%	25%	746
Harvard University[b]	1.43	20%	37%	24%	20%	51
University of Michigan: Ann Arbor	1.18	31%	33%	23%	14%	1267
Web-based studies[c]	1.10	39%	25%	22%	13%	525
Bowling Green University	0.87	50%	25%	13%	12%	52
University of Michigan: Dearborn	0.83	51%	22%	21%	6%	154
Michigan State University	0.79	49%	29%	16%	6%	118
University of Toledo	0.57	64%	21%	10%	5%	138
Overall	1.24	33%	28%	23%	17%	3428

Die Fragestellungen sowohl der SRIS von Silvia (2021) als auch des CRT von Frederick (2005) wurden aus der Originalsprache ins Deutsche übersetzt. Dabei wurde darauf geachtet, dass inhaltlich wie sprachlich eine große Nähe zur Originalformulierung beibehalten wurde, um Übersetzungsverzerrungen zu reduzieren. Zur zusätzlichen Absicherung wurde die Methode der Rückübersetzung angewendet (Ali, 2016; Brislin, 1970, 1976), da Wortbedeutungen häufig nicht gleichbedeutend übersetzt werden und andere Bedeutungsnuancen enthalten können (Mesquita et al., 1997). Nach erfolgter Übersetzung wurden diese übersetzten Items von einer bilingualen Person zurückübersetzt, um etwaige semantische Unterschiede aufzudecken, und in notwendigen Fällen Anpassungen vorgenommen.

Die eingesetzten Fragebögen wurden vor ihrem Ersteinsatz pilotiert, um die Tauglichkeit und Qualität sicherzustellen (Riese & Reinhold, 2014). Dazu fand hauptsächlich ein Ersteinsatz durch Teilnehmende außerhalb der eigentlichen Stichprobe und Forschende der Chemiedidaktik statt, welchem eine kommunikative Validierung folgte, um etwaige Herausforderungen beim Verständnis der Formulierung oder Ausrichtung der Items feststellen und diese verbessern zu können (Krüger et al., 2014; Reinders & Bergs-Winkels, 2022).

LimeSurvey wurde so programmiert, dass jede Frage auf einer eigenen Seite dargestellt wurde, um Ablenkungen, Konsistenz- und Ankereffekte zu reduzieren (Chapman & Johnson, 2002; Moosbrugger & Brandt, 2020b; Salancik & Pfeffer, 1977). Gleichzeitig wurden innerhalb einzelner Items, die aus mehreren Teilen bestanden, die Reihenfolge der einzelnen Teile randomisiert, um Reihenfolgeeffekte,

wie den Response-Bias, zu verringern (Krosnick et al., 2014; Moosbrugger & Brandt, 2020b). Zur Beantwortung der Fragen hatten die Teilnehmenden kein Zeitlimit, um keinen Zeitdruck zu erzeugen. Die meisten Teilnehmenden benötigten einen Zeitraum von etwa 20 bis 25 Minuten zur Beantwortung der Fragebögen. Die finale Version des verwendeten Fragebogens ist im Anhang (A.1.2) dokumentiert.

4.2.4 Interviewerhebungen

Zur Sicherstellung der Validität der Datenerhebung werden in Anlehnung an Niebert und Gropengießer (2014) die Gütekriterien nach Mayring (2010) angelegt. Daher werden nachfolgend zur besseren Nachvollziehbarkeit die Vorgänge bei der Datenerhebung dokumentiert und die Bedingungen transparent gemacht, unter denen eine Weiterverarbeitung der Daten stattfand.

Die Interviewerhebungen kamen ab der zweiten Studie S2 zum Einsatz, da sich Interviews als offene Befragungsmethode besonders zur Exploration sowie zum Erfassen subjektiver Bedeutungsstrukturen und individueller Sichtweisen eignen (Reinders, 2022b). Das Interview stellte im Kontext dieser Zielsetzung in seiner asymmetrischen Kommunikationsform eine systematische Erhebungsmethode dar, die trotz ihrer Struktur ein hohes Maß an Flexibilität zur dynamischen Anpassung zulässt, um spezifische Antworten der befragten Person weiter zu erkunden (Christel, 2019; Lamnek & Krell, 2016; Przyborski & Wohlrab-Sahr, 2021; Reinders, 2022b). Da ein ähnlicher Betrachtungskontext wie bei Kolstø (2001) vorlag, welcher Interviews mit SuS erprobte und durchführte, um Einblicke in deren Einschätzung von Informationen hinsichtlich ihrer Glaubwürdigkeit zu erhalten, wurde sich zur Fokussierung auf das Erkenntnisinteresse und zur Aufrechterhaltung einer groben Struktur innerhalb der Interviews ebenfalls dazu entschieden, semistrukturierte Interviewleitfäden für die Befragungen einzusetzen (Christel, 2019; Niebert & Gropengießer, 2014; Przyborski & Wohlrab-Sahr, 2021). Diese geringere Standardisierung ermutigt dabei zu lebensnäheren Antworten, indem eine freie Antwortformulierung in der eigenen Sprache ermöglicht wird. Hierfür wurde ein weicher Kommunikationsstil gewählt, um eine Offenheit und nicht bewertende Haltung zu vermitteln, und nur bei einem Themenwechsel aktiv eingegriffen (Koolwijk, 1974; Lamnek & Krell, 2016). Zur Erzeugung einer ungezwungenen Gesprächsatmosphäre, Eingehen einer Kommunikationsbeziehung sowie zur Reduktion eines Gefühls von Leistungsdruck wurde das *Du* als Anrede gewählt (Hussy et al., 2010; Lamnek & Krell, 2016). Die Gesprächsführung wurde flexibel gestaltet, um einen möglichst natürlichen Gesprächsfluss zu erzielen. Denkpausen der Teilnehmenden wurden toleriert und sich

vorwiegend im Hintergrund gehalten, um den Redeanteil der Befragten hochzuhalten (Niebert & Gropengießer, 2014). Des Weiteren wurde versucht, Suggestivfragen und Belehrungen zu vermeiden sowie die Fragen klar und kurz zu halten, um das Verständnis zu erleichtern (Reinders, 2022b).

Rekurrierend auf das Erkenntnisinteresse, Kritisches Denken in komplexen Bewertungssituationen erfassbar zu machen, wurde sich in Anlehnung an das von Sadler et al. (2007) entwickelte *Socioscientific Reasoning* (SSR) (vgl. Kap. 2.1.6) dazu entschieden, SSI als Bezugsrahmen in den Interviews zu verwenden. Diese komplexen Kontexte bieten aufgrund ihrer kontroversen Natur und ihrer Uneindeutigkeit ideale Bedingungen zur Untersuchung Kritischen Denkens, da sie einen ebenso komplexen wie multiperspektivischen Betrachtungsansatz erfordern (Bybee, 1997; Gunn et al., 2008; Kolstø et al., 2006; López-Fernández et al., 2022; Osborne, 2014; Rafolt et al., 2019; Yacoubian, 2015; Zeidler, 2014). Gleichzeitig ist anzunehmen, dass Befragte bei der Bewertung von Kontexten mit einer eindeutigen Lösung deutlich früher zu einer abschließenden Antwort kommen, was die Ergiebigkeit des Interviews reduzieren würde.

Die Kontextualisierung erfolgte durch dichotom formulierte Fragestellungen, die zwischen zwei alternativen Möglichkeiten der Positionierung unterschieden und als Stimulus für eine vertiefende Auseinandersetzung mit der Thematik dienten. Zielsetzung war das Anregen der Teilnehmenden zum Abwägen der Argumente beider Seiten und zur Reflexion der Konsequenzen einer klaren Positionierung aus ethischer, moralischer, sozialer und ökologischer Perspektive sowie damit einhergehend ein besserer Einblick in die Denk- und Bewertungsstrukturen der Teilnehmenden. Methodologisch bot die dichotome Struktur zugleich eine bessere Vergleichbarkeit der Ergebnisse, da alle Teilnehmenden mit identischen Entscheidungssituationen konfrontiert wurden. Dies minimierte mögliche Verzerrungen durch unterschiedlich wahrgenommene Optionen und ermöglichte eine klarere Zuordnung und Kategorisierung der Ergebnisse. Zusätzlich spiegelte diese ambivalente Struktur reale Entscheidungssituationen wider, wie sie im alltäglichen Leben, beispielsweise bei politischen Wahlen oder ethischen Dilemmata, auftreten. Dadurch wurde ein praxisnaher Rahmen geschaffen, der sowohl die Relevanz der Entscheidungen als auch die Übertragbarkeit der Ergebnisse auf reale Situationen erhöhte. Die Kontroversität und oft unklare Evidenzlage, die Socioscientific Issues (SSI) per definitionem auszeichnen, wurden durch die starke Polarisierung der Fragestellungen hervorgehoben und unterstrichen die realweltliche Herausforderung, zwischen Alternativen zu wählen, die

keine objektiv beste Lösung bieten, sondern stark von subjektiven Werten und Präferenzen beeinflusst werden. Da diese subjektiven Werte und Präferenzen im Zentrum des Erkenntnisinteresses standen, erwies sich die dichotome Struktur als besonders geeignet, um die gewünschten Erkenntnisse zu generieren.

Je nach Studie wurden verschiedene Kontexte verwendet (vgl. Kap. 4.2.6), in denen eine Entscheidungssituation als Setting suggeriert wurde. Um den gesellschaftlichen Bezug zu verstärken, wurde eine Situation angenommen, in der jede der beiden Alternativen für die Position einer Partei bei einer anstehenden Wahl oder Positionierung in einem Austausch mit Freunden stand. Diese partizipatorische Komponente sollte die Auswirkungen der Entscheidung hervorheben und eine realistische Lebensweltsituation darstellen. Gleichzeitig kann damit dem Gütekriterium der Angemessenheit des Erhebungsinstruments entsprochen werden (Steinke, 2019), indem sich möglichst nah an tatsächlichen lebensweltlichen Fragestellungen orientiert wurde und der Gesprächsansatz die natürlichste Form der Kommunikation darstellt.

Aufgrund der Erhebung von Bewertungen und Bewertungsstrategien sowie der Verwendung einer komplexen Kontextualisierung kann das Leitfadeninterview als Dilemma-Interview präzisiert werden (Christel, 2019; Colby & Kohlberg, 1987; Niebert & Gropengießer, 2014). In diesem sollten aber nicht nur Sichtweisen erhoben werden, sondern auch Herangehensweisen und Denkperspektiven. Um der angestrebten Gelingensbedingung (vgl. Kap. 4.1) nachzukommen, die Perspektiven des Denkens und nicht das (Fach-)Wissen zu erheben, wurde die Ausrichtung dahingehend verändert, dass den Teilnehmenden erlaubt war, Fragen zu formulieren, die sie sich im jeweiligen Bewertungskontext stellten. Diese Vorgehensweise wurde unter der Annahme erprobt, dass die Fähigkeit zu denken durch die Fähigkeit determiniert wird, sich selbst Fragen zu stellen (Clough, 2008; Dewey, 1933; Paul & Elder, 2014). So konnten die Kontexte auch dann Verwendung finden, wenn seitens der Teilnehmenden kein spezifisches Fachwissen vorhanden war.

Die Interviews wurden schwerpunktmäßig in Präsenz geführt und die Fragen in mündlicher Form gestellt. Nur in drei Fällen musste auf die Online-Konferenz-Plattform *Zoom* (Zoom Video Communications, Inc., 2024) zurückgegriffen werden, welche an der Universität zu Köln im Rahmen der digitalen Lehre und auch für offizielle digitale Veranstaltungen eingesetzt wird. In diesen Fällen wurden die Interviews audiovisuell geführt, um auch trotz der Asymmetrie des Interviewformates nonverbal den weichen Kommunikationsstil beibehalten zu können. Die Gespräche wurden in Präsenz durch Aufnahmegeräte mit lokalem Datenspeicher festgehalten und unter Berücksichtigung des Datenschutzes weiterverarbeitet (vgl. Kap. 4.2.1). Die digital

geführten Interviews wurden in Zoom mithilfe der internen Aufnahmefunktion aufgezeichnet und allein die Tonspur weiterverwendet.

Der Leitfaden zum Interview wurde tabellarisch erstellt und kann im Anhang (A.1.1) eingesehen werden. Diese Aufbereitung diente zum einen dem schnellen Überblick, um nicht im Gespräch abzulenken, zum anderen zur Wahrung der logischen Struktur (Niebert & Gropengießer, 2014). Ad-hoc-Interventionen wurden durchgeführt, wenn dies im Kontext gewinnbringend und notwendig erschien, und sehr bewusst eingesetzt, um keine Artefakte zu erzeugen (ebd.)

Den Anfang des Interviews bildete die Frage nach der Einwilligung zur Aufnahme des Gesprächs seitens der Teilnehmenden, um das Einverständnis sicherzustellen und auf den Beginn der Aufnahme hinzuweisen. Daran anschließend folgte die Einstiegsphase, in der ein lockerer Beginn in das Gespräch erfolgte, um den wahrgenommenen Druck bei den Teilnehmenden zu reduzieren (Reinders, 2022b). Im Weiteren wurde dann der Rahmen des Interviews erläutert und über den Datenschutz aufgeklärt sowie, falls nötig, ein Code generiert. Insgesamt war es das Ziel, die Angst vor dem Interviewsetting zu nehmen und deutlich zu machen, dass es sich um keine Prüfungssituation handelt und alle Antworten pseudonymisiert weiterverarbeitet werden. In diesem Rahmen wurde auch auf das Recht an den eigenen Daten hingewiesen und der Interviewablauf erläutert, um in den Modus des Gesprächs einzuführen (ebd.) und Vorbehalte durch die asymmetrische Gesprächsführung zu reduzieren.

Auf die Einstiegsphase folgte die Aufwärmphase, um das Gespräch in Gang zu setzen. Dies fand durch eine Eisbrecher-Frage statt, welche noch keine hohe Relevanz für das Interview hatte und leicht zu beantworten war (Döring & Bortz, 2016).

Zur Erstellung der Hauptphase des Leitfadens wurden zunächst erste Fragekategorien erstellt und Fragen formuliert, wobei es das Ziel war, vom Allgemeinen zum Spezifischen zu verfahren, um kurze Antworten und ein Hemmen der Gesprächsführung zu vermeiden (Przyborski & Wohlrab-Sahr, 2021). Nach der generellen Frage nach relevanten Aspekten im Entscheidungsprozess zum ausgewählten Setting sollten so die Nachfragen immer konkreter werden und ein Priming durch eine direkt explizite Formulierung im Vorfeld umgangen werden. Die Fragen verliefen so von Nachfragen zum Vorgehen im Entscheidungsprozess über das tatsächlich vorhandene Fachwissen bis hin zur Reflexion interner oder externer Beeinflussungsfaktoren. Den Abschluss bildeten explizite Verständnisfragen. Die Hauptfragen im Leit-

faden dienten dabei als offener Stimulus für freie Ausführungen der Befragten. Zusätzlich zur Hauptfrage einer jeden Kategorie wurden spezifischere Unterfragen formuliert, die dazu dienen sollten, relevante Informationen einzuholen, sollten diese nicht von selbst auf die Initialfrage folgen (ebd.). Es folgt ein Beispiel aus dem Abschnitt

Fragenformulierung 1:

Hauptfrage:

- Welche Fragen stellst du dir, wenn du vor so eine Entscheidung gestellt wirst?

Unterfragen:

- Was sind für dich hier relevante Aspekte?
- Welche Fragen hättest du gerne beantwortet, um diese Entscheidung treffen zu können?

Tabelle 10 zeigt den immer konkreter werdenden Charakter der Fragen, von der freien Formulierung relevanter Fragen im Entscheidungsprozess über Stimuli bis hin zur expliziten Nachfrage über das Verständnis der Begrifflichkeiten.

Tabelle 10: Schematische Darstellung des Interviewleitfadens von Studie S2

Fragestellung	Anmerkung
1. Einleitung und Vorstellung	
2. Einstieg	
3. Entscheidungsfindung **4. Fragenformulierung 1** **5. Fragenformulierung 2**	Ungerichteter Einstieg in Bewertungssituation, wenig explizit
6. Reflexionsstimulus kognitiv **7. Nachfrage zur Reflexion**	Fokussierung der ***kognitiven*** Ebene
8. Reflexionsstimulus affektiv **9. Explizite Reflexion**	Fokussierung der ***affektiven*** Ebene (weitere Beeinflussungsfaktoren)
10. Selbsteinschätzung und Verständnis	Explizite Thematisierung der Begrifflichkeiten
11. Abschluss	

Während der Hauptphase des Interviews wurden die Teilnehmenden gebeten, ihre Gedanken während der Positionierung innerhalb des Kontextes periaktional zu verbalisieren. Diese Art des lauten Denkens diente der Erfassung kognitiver Prozesse, um so die grundlegenden Gedankengänge im Bewertungsprozess nachvollziehen zu können (Hussy et al., 2010; Sandmann, 2014). Dabei muss jedoch hervorgehoben

werden, dass es sich nicht um die Anwendung des lauten Denkens als Forschungsmethode handelte, sondern um einen unterstützenden Einblick in die Denkprozesse während der Interviews und damit um das laute Denken in seiner Ursprungsdefinition (Ericsson & Simon, 1993; Huber & Mandl, 1994; Knoblich & Ollinger, 2006; Sandmann, 2014).

Den Abschluss bildete die Ausstiegsphase, welche eingeleitet wurde, wenn ersichtlich war, dass sich die Behandlung der Fragen erschöpfte (Reinders, 2022b). Hier wurde abschließend nach Ergänzungen gefragt und die Möglichkeit eröffnet, selbst noch Fragen zu stellen, sollten welche offen sein. Nach dem Beenden der Aufnahme ergab sich häufig noch eine Nach-Interviewphase, in der durch die ohne Dokumentation entstehende Auflockerung zum Teil weitere interessante Gesprächsverläufe zustande kamen und die interviewte Person wieder aus dem Interviewsetting hinausgeführt wurde (ebd.)

Der finale Interviewleitfaden wurde mit $N = 3$ Studierenden des Lehramts eines anderen naturwissenschaftlichen Faches pilotiert (vgl. Kap. 4.2.2). Auch hier kam das laute Denken zur Überprüfung der Fragengüte im Pilotierungsprozess zum Einsatz (Sandmann, 2014). Neben dem geplanten periaktionalen lauten Denken wurde zur Pilotierung das postaktionale oder auch retrospektive laute Denken angewandt, um Verständnisprobleme oder Herausforderungen im Interview transparent zu machen (Hussy et al., 2010; Sandmann, 2014). Auf Basis der Einsichten aus den Pilotierungen wurden Anpassungen an einzelnen Fragenformulierungen oder Ergänzungen vorgenommen, wobei die Fragen inhaltlich aber im Leitfaden verblieben. Ein Beispiel stellt die Ausschärfung der Fragestellungen dar, um Unsicherheiten aufgrund ungerichteter und abstrakter Formulierungen zu vermeiden. So führte die ursprüngliche ungerichtete Frage *Gibt es Fragen, die du dir jetzt stellst?* in den Pilotierungen verstärkt zu Unsicherheiten und war potenziell durch eine einfache Ja- oder Nein-Antwort seitens der Teilnehmenden erschöpft. Darauf reagierend wurde die Frage umformuliert in *Welche Fragen stellst du dir, wenn du vor so eine Entscheidung gestellt wirst?*, um einen offeneren Raum zu schaffen, in dem gezielt Fragen eingebracht werden sollen. Gleichzeitig verankert diese Formulierung die Teilnehmenden stärker im Kontext, was eine klarere Einordnung ermöglicht. Um dennoch auf mögliche Unsicherheiten seitens der Teilnehmenden zu reagieren, wurden ergänzende Unterfragen formuliert. So wurde etwa die Frage *Welche Fragen hättest du gerne beantwortet, um diese Entscheidung treffen zu können?* hinzugefügt, da diese kleine Paraphrasierung der Ursprungsfrage den Fokus von der Selbstwahrnehmung (Welche Fragen stelle ich mir?) auf die Situation

selbst (Welche Fragen müssen beantwortet werden?) lenkt. Diese Erweiterung berücksichtigt auch, dass nicht alle Menschen, sondern nur schätzungsweise 25 % in ihren Denkprozessen eine innere Stimme verwenden (Hurlburt et al., 2013). Diese zusätzliche Frage bietet daher eine alternative Perspektive, die den Teilnehmenden eine bessere Orientierung und Reflexion erlaubt, ohne jedoch den Fokus der Fragengenese und des Erkenntnisinteresses zu verlieren.

Zwischen Studie S2 und S3 wurde der Interviewleitfaden farblich angepasst, um eine Zuordnung der Interviewteile zu den Dimensionen des parallel entwickelten *Triadischen Modells Kritischen Denkens in den Naturwissenschaftsdidaktiken* (s. Abbildung 13; vgl. Kap. 3.5) zu ermöglichen. Die Fragen des Leitfadens blieben inhaltlich unverändert, wurden jedoch leicht neu gruppiert. Da die Anpassung vorrangig der Anwenderfreundlichkeit diente und ein semistrukturiertes Leitfadeninterview keine starre Struktur erfordert, konnte die Vergleichbarkeit zwischen den Stichproben weiterhin gewährleistet werden.

Tabelle 11: Schematische Darstellung des Interviewleitfadens ab Studie S3

Fragestellung	Anmerkung
1. Einleitung und Vorstellung	
2. Einstieg	
3. Problemaufriss **4. Fragenformulierung 1** **5. Fragenformulierung 2**	Ungerichteter Einstieg in Bewertungssituation, wenig explizit
6. Vorgehensbeschreibung **7. Wissensreflexion**	Fokussierung der *kognitiven* Ebene
8. Selbstreflexive Aspekte (implizit) **9. Selbstreflexive Aspekte (explizit)**	Fokussierung der *affektiven* Ebene
10. Selbsteinschätzung und Verständnis	Explizite Thematisierung der Begrifflichkeiten
11. Abschluss	

Auch hier wird anhand der Tabelle 11 der graduelle Übergang vom Allgemeinen zum Spezifischen deutlich, jedoch zusätzlich im Kontext der einzelnen Dimensionen des Modells. Der zugehörige Interviewleitfaden kann im Anhang (A.1.1) eingesehen werden.

4.2.5 Vermittlungseinheit

Ziel der Vermittlungseinheit war nicht primär die Entwicklung einer Fördermöglichkeit, sondern die Erprobung des entwickelten Instruments in einem Pre-Post-Follow-Up-Testformat. Auf diese Weise konnte untersucht werden, ob das entwickelte Instrument die Dimensionen misst, welche durch die Vermittlungseinheiten angesprochen und gefördert werden sollten. Die Fördermöglichkeiten wurden im Vorfeld der Ersterhebungen entwickelt und in ein bestehendes Seminarkonzept der Universität zu Köln (*Seminar zu fachbezogenen Lern- und Kommunikationsprozessen*, *B5 HR/GG B3 SP)* integriert. Im Sommersemester 2022 erfolgte die erste Erprobung innerhalb dieses Seminars. Nach diesem Erstdurchlauf wurden die Fördermöglichkeiten im Rahmen des Forschungsprojektes im Sommersemester 2023 im gleichen Seminarformat angewendet. Dieses bestand aus vier Kernthemen (Sitzung 1-4; Kompetenzen, Unterrichtsmethoden, Individuelle Förderung und Medien im Chemieunterricht) und konnte im Rahmen forschungsbasierter Lehre um Inhalte mit Bezug zum Dissertationsprojekt (Sitzungen 5 bis 12) ergänzt werden (Krüger et al., 2014). Die letzte Sitzung diente als Rückblick und Reflexion des Seminars (Tabelle 12).

Der forschungsrelevante Anteil des Seminars, die Sitzungen fünf bis zwölf, setzte sich aus dreimal zwei Einheiten zu verschiedenen Bereichen der Bewertung zusammen, wobei jede Einheit eine Dauer von 1,5 Stunden hatte.

Die Sitzungen fünf und sechs dienten zur Erweiterung des Verständnisses naturwissenschaftlicher Arbeits- und Denkweisen durch eine exemplarische Vermittlung von Nature of Science (NOS) im Rahmen eines Workshops auf Basis der Forschung von Müller (2021). Dieser bestand aus einer Reihe von NOS-assoziierten Aktivitäten, welche unter anderem die Vorläufigkeit und die soziokulturelle Eingebundenheit naturwissenschaftlicher Erkenntnisse in den Fokus stellten, aber auch die Subjektivität und Kreativität in den Naturwissenschaften hervorheben sollten. Diese Eigenschaften und Teilbereiche der Lederman Seven (Lederman, 1992) stellen dabei besonders schwer greifbare Merkmale naturwissenschaftlicher Forschung dar, welche häufig von Fehlvorstellungen begleitet werden. Zusätzlich wurden die von McComas (1998) formulierten Mythen über NOS diskutiert, die Argumente für die Vermittlung von NOS (Driver et al., 1996; Thomas & Durant, 1987) erörtert wie auch die Delphi-Studie von Osborne et al. (2003) angeführt, um ein umfassenderes Bild von NOS, ihrer Definition sowie der Grenzen und Möglichkeiten zu zeichnen (vgl. Kap. 2.1.2).

Darauf aufbauend wurde in Sitzung sieben das Abgrenzungsproblem zwischen Wissenschaft und Pseudowissenschaften erörtert (vgl. Kap. 2.1.5) (Fasce, 2017; Hansson, 2013; Popper, 1976) sowie Möglichkeiten, Informationen einschätzen zu können, vorgestellt (Mahner, 2013). In Sitzung acht wurde daraufhin die Brücke zur Wissenschaftskommunikation und der Rezeption naturwissenschaftlicher Erkenntnisse in der Gesellschaft geschlagen. In diesem Zusammenhang wurden vor allem die Auswirkungen von Fehlinformationen, Falschinformationen und Desinformationen auf die Gesellschaft thematisiert (Brainard et al., 2020; Brainard & Hunter, 2020; Feierabend et al., 2022; Meßmer et al., 2021; Vodafone Stiftung Deutschland gGmbH, 2020). Zusätzlich wurden die Möglichkeiten und Herausforderungen diskutiert, die sich aus den Auswirkungen sozialer Medien auf die Wissenschaftskommunikation ergeben (vgl. Kap. 2.1.3) (Archila et al., 2021; Höttecke & Allchin, 2020).

In Sitzung 9 wurde dann erstmals der Begriff der kognitiven Verzerrungen eingeführt und ihre Bedeutung für die Bewertung von naturwissenschaftlichen Kontexten (SSI) vermittelt (Gigerenzer, 1991; Gilovich et al., 2002; Kahneman et al., 1982; Pronin et al., 2002; Thompson et al., 2013). In diesem Zuge wurde auch die Dual-Process Theory aufgegriffen sowie für ihre Implikationen sensibilisiert (vgl. Kap. 2.3.2) (Evans, Jonathan St. B. T., 2006; Gilovich et al., 2002; Kryjevskaia et al., 2021; Osman, 2004).

In Sitzung 10 folgte die konkrete Definition des Konstruktes des Kritischen Denkens (Abrami et al., 2015; Bailin et al., 1999b; Brambatti Guzzo & Dall'Alba, 2017; Byrne & Johnstone, 1987; Ennis, 1989; Facione, 1990; Halpern, 2014; Kogut, 1996; Minott et al., 2019; Paul & Elder, 2014; Rafolt et al., 2019) sowie die Definition und Bedeutung der Selbstreflexion innerhalb des Kritischen Denkens (vgl. Kap. 3.3). Der systemische Betrachtungsansatz, bestehend aus epistemologischem Wissen, naturwissenschaftlichem Wissen und selbstreflexivem Denken, wurde vermittelt (vgl. Kap. 3.5) und die eigens dafür entwickelte Methode der *Systemisch-reflexiven Stellungnahme* (SrS) angebahnt.

Tabelle 12: Übersicht über die Seminarsitzungen im SoSe23, Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP)

Sitzung	Inhalt
1	Begrüßung, Organisatorisches und Einführung Pre-Test-Fragebogen Einblick in die Kompetenzbereiche und Ziele der Chemiedidaktik
2	Unterrichtsmethoden
3	Individuelle Förderung: Differenzierung und Diagnostik im (inklusiven) Chemieunterricht
4	Umgang mit Medien im Chemieunterricht: Experimente, Modelle, didaktische Programme und mehr
5	Nature of Science 1 – Workshop; angeleitete Selbsterfahrung mit NOS
6	Nature of Science 2 – Diskussion und Reflexion von NOS
7	Science and Pseudoscience – Das Abgrenzungsproblem und Annäherungen an Differenzierungsmöglichkeiten
8	Medien und Naturwissenschaften – Wissenschaftskommunikation in Zeiten digitaler Medien; Einschätzung von Quellen
9	Kognitive Verzerrungen – Was unser Denken und Verhalten beeinflusst; Dual-Process Theory
10	Kritisches Denken – Begriffsdefinition und Relevanz bei der Bewertung naturwissenschaftlicher Kontexte
11	Systemisch-Reflexive-Stellungnahme – Vierschrittmethode zur Förderung der Multiperspektivität in Bewertungskontexten anhand exemplarischer SSI
12	Weiterführung und Reflexion der Systemisch-reflexiven Stellungnahme
13	– Post-Test-Fragebogen – Reflexion des Seminars – Evaluierung, Ausblick und Abschluss

Systemisch-reflexive Stellungnahme

Die *Systemisch-reflexive Stellungnahme* (SrS) stellt eine angeleitete, konkrete Bewertungssituation dar und besteht aus vier aufeinanderfolgenden Schritten mit dem Ziel, sowohl das naturwissenschaftliche Wissen, das epistemologische Wissen als auch die Selbstreflexion der affektiven Einflüsse in einer systemischen Weise in verschiedenen Kontexten zu reflektieren (vgl. Kap. 3.5).

Die Lernenden konnten einen von fünf vorgegebenen Kontexten auswählen und behielten diesen während des gesamten Prozesses bei. Alle wählbaren Kontexte stellten dabei sogenannte Socioscientific Issues (SSI) dar, komplexe gesellschaftliche Problemstellungen mit einem naturwissenschaftlichen Bezug ohne eindeutige Lösung (vgl. Kap.2.1.6) (Zeidler, 2014; Zeidler & Kahn, 2014; Zeidler & Nichols, 2009),

welche aufgrund ihrer kontroversen Natur einen multiperspektivischen Betrachtungsansatz erfordern und daher hier als konkrete Anforderungssituationen dienten, um Fähigkeiten, Kompetenzen und Strategien anzuwenden (Bruder et al., 2017; Klieme et al., 2003; Lersch, 2007, 2022).

Die auswählbaren Kontexte wurden als dichotome Entscheidung zwischen zwei Extremen vorformuliert und Rechercheansätze inhaltlich impliziert, um die Hemmschwelle bei der initialen Auseinandersetzung zur reduzieren (Tabelle 13).

Tabelle 13: Auswählbare Kontexte für die Systemisch-reflexive Stellungnahme

Kontext 1: Fracking (Hydraulic Fracturing)	Fragestellung: Sollte die Praxis des Frackings, also die hydraulische Frakturierung von Gesteinsschichten zur Gewinnung von Erdgas oder Erdöl, fortgesetzt werden, um Energievorräte zu erschließen und die wirtschaftliche Entwicklung zu fördern, oder sollten wir aufgrund möglicher Umweltauswirkungen und Risiken den Einsatz von Fracking einschränken oder verbieten?
Kontext 2: Plastik	Fragestellung: Sollten Einwegplastikprodukte, wie Plastiktüten und Einwegflaschen, stärker reguliert oder verboten werden, um die Auswirkungen auf die Umwelt durch Plastikmüll zu reduzieren, oder sollten wir den Schwerpunkt eher auf verbessertes Recycling und die Entwicklung biologisch abbaubarer Kunststoffe legen, um die Vorteile von Plastikprodukten beizubehalten?
Kontext 3: Nanotechnologie	Fragestellung: Sollte die Entwicklung und Anwendung von Nanotechnologie strenger reguliert werden, um potenzielle Risiken für die Gesundheit und Umwelt zu minimieren, oder sollten wir die Chancen und Potenziale der Nanotechnologie nutzen und die Regulierung flexibel gestalten, um Innovationen und wirtschaftlichen Fortschritt zu ermöglichen?
Kontext 4: Chemie in der Landwirtschaft	Fragestellung: Sollten chemische Pestizide und Düngemittel in der Landwirtschaft weiterhin weit verbreitet eingesetzt werden, um Ernteerträge zu maximieren und die Welternährungssicherheit zu gewährleisten, oder sollten wir verstärkt auf alternative Methoden wie biologischen Anbau und nachhaltige Landwirtschaft setzen, um Umweltauswirkungen und potenzielle Gesundheitsrisiken zu minimieren?
Kontext 5: Atomenergie	Fragestellung: Sollte Atomenergie als eine bedeutende Energiequelle in der Zukunft weiter ausgebaut werden, um den steigenden Energiebedarf zu decken und den Klimawandel zu bekämpfen, oder sollten wir verstärkt auf erneuerbare Energien setzen und Atomenergie aufgrund ihrer potenziellen Risiken und Auswirkungen einschränken oder sogar auslaufen lassen?

Nach selbstständiger Auswahl eines Kontextes, bei dem die Studierenden interessengeleitet entscheiden konnten, um eine Grundmotivation zu erzeugen, erfolgte der erste Schritt der SrS:

Schritt 1: Sachanalyse (Einzelarbeit)

Im ersten Schritt der Systemisch-reflexiven Stellungnahme führten die Studierenden in Einzelarbeit eine (Web-)Recherche und Sachanalyse des ausgewählten Kontextes sowie eine Bewertung der dabei verwendeten Quellen durch. Der Fokus dieses Schrittes lag auf dem Wissenserwerb sowie der Einschätzung der recherchierten Informationen. Hier galt es, die Informationen zu sammeln, die auf rationaler Ebene einer Stellungnahme als Fundament dienen können.

Schritt 2: Dialogisches Selbstreflexionstraining (Partnerarbeit)

Der zweite Schritt fand in Partnerarbeit statt und fokussierte die Selbstreflexion, wobei sich die Partner*innen gegenseitig unterstützten. Beginnend mit einem kurzen Statement zum aktuellen Stand der eigenen Position wurden im Anschluss anhand zweier unterschiedlicher Impulsleitfäden gegenseitig abwechselnd reflexive Fragen gestellt und so der Prozess der Erkenntnisgewinnung und Positionierung hinterfragt. Die/Der Fragenstellende hatte dabei immer die Möglichkeit, ergänzende Fragen zu stellen, die dem reflexiven Prozess dienlich sein könnten. Erkenntnisse aus diesem Prozess wurden ebenfalls durch die Studierenden dokumentiert und Veränderungen der eigenen Position deutlich gemacht. Hier sollten die eigene affektive Ebene beleuchtet und etwaige Einflüsse auf die Entscheidung seitens System 1 hinterfragt werden.

Schritt 3: Systemisch-Reflexive Stellungnahme (Gruppenarbeit)

Im dritten Schritt fanden sich die Studierenden in Gruppen von vier bis sechs Personen zusammen. Jede*r Lernende nahm dann in einem kurzen Vortrag Stellung zu dem bearbeiteten Kontext mit dem Ziel, die Erkenntnisse aus der Sachanalyse und dem selbstreflexiven Teil miteinander in Verbindung zu bringen und zu verdeutlichen. Es sollte darauf eingegangen werden, wie das Wissen in der Sachanalyse gewonnen wurde, welche Quellen verwendet wurden und inwieweit eigene Prägungen bei der Positionierung eine Rolle spielten. Im Anschluss an die Stellungnahme der Einzelperson folgte jeweils eine kurze Diskussionsrunde, in der jedes Gruppenmitglied Fragen zum Ablauf und zur Position stellen konnte, mit dem Ziel, eine differenzierte Betrachtung zu ermöglichen.

Schritt 4: Prozessreflexion (Einzelarbeit)

Der vierte und letzte Schritt fand wieder in Einzelarbeit statt und hielt die Studierenden dazu an, ihre gemachten Erfahrungen in diesem Entscheidungsprozess zu reflektieren, Herausforderungen zu diskutieren und zu erläutern, welche Verzerrungen

bei ihnen selbst spürbar waren. Dieser letzte Teil diente zum einen der Retrospektion des Prozesses, zum anderen der Introspektion und Selbsterkenntnis, wie die eigene Position zustande kam, wie sie sich veränderte und was bei zukünftigen Entscheidungen stärker fokussiert werden sollte.

Ein Überblick über die konkreten Arbeitsanweisungen ist dem Anhang (A.1.3) beigefügt. Die Arbeitsanweisungen unterschieden sich in Abhängigkeit vom enthaltenen Impulsleitfaden, welche inhaltlich different waren. Dies lag darin begründet, dass eine Doppelung der Fragen keinen effektiven Mehrwert geboten hätte, weil beim Stellen einer Frage bereits ein damit verknüpfter Denkprozess abläuft. So wurde jede Frage im Tandem nur einmal thematisiert, was die Gespräche abwechslungsreicher gestalten sollte. Der zweite Teil des Impulsleitfadens kann im Anhang (A.1.3.2) eingesehen werden. Die anderen Aufgabenstellungen blieben jedoch deckungsgleich.

In Studie S5 wurde die Vermittlungseinheit ebenfalls durchgeführt, aber aufgrund des Umstands, dass es sich um das Seminar eines anderen Dozenten handelte, stark komprimiert (Tabelle 14).

Tabelle 14: Übersicht über die Seminarinhalte in Studie S5

Sitzung	Inhalt
1	Begrüßung und Einführung Nature of Science – Workshop, Diskussion und Reflexion von NOS
2	Science, Pseudoscience und Medien – Wissenschaftskommunikation und ihre Herausforderungen
3	Kritisches Denken und der Einfluss interner und externer Faktoren
4	Systemisch-Reflexive-Stellungnahme – Vierschrittmethode zur Förderung der Multiperspektivität in Bewertungskontexten anhand exemplarischer SSI
5	Post-Test-Fragebogen Reflexion und Verabschiedung

Die Inhalte der Einheiten waren dennoch vergleichbar, auch wenn weniger Raum für eine diskursive Auseinandersetzung und Reflexion im Plenum gegeben war.

Die Aufgabenstellung der Systemisch-reflexiven Stellungnahme unterschied sich in Studie S5 im Vergleich zu Studie S3 einzig in den zur Wahl gestellten Kontexten, da eine Überschneidung mit den Interviewkontexten verhindert werden sollte. Eine Übersicht über die veränderten zur Verfügung gestellten Kontexte der Aufgabenstellung 1 ist im Anhang (A.1.3.3) zu finden.

4.2.6 Verwendete Kontexte

Das Ziel des Interviews war es, die für die Proband*innen relevanten Aspekte bei gesellschaftlichen Problemstellungen mit naturwissenschaftlichem Bezug zu ergründen und ihre Denk- und Bewertungsprozesse im Rahmen der Untersuchung der Fähigkeit zum Kritischen Denken offenzulegen. Da sich komplexe Denkfähigkeiten nur dann zeigen können, wenn auch komplexe Anforderungssituationen vorhanden sind (Lersch & Schreder, 2013), wurde sich in Anlehnung an den SSR-Ansatz von Sadler et al. (2007) dazu entschieden, SSI als Kontexte zu verwenden (vgl. Kap. 2.1.6). Diese zeichnen sich vor allem durch ihre Uneindeutigkeit und kontroverse Natur aus (Chen & Xiao, 2021; Sadler, 2011; Zeidler et al., 2019). Besonders die moralischen und ethischen Herausforderungen erfordert dabei metakognitive Denkprozesse (Mestre & Cocking, 2002; Zeidler & Sadler, 2007) und Kritischem Denken kommt bei der Auseinandersetzung mit SSI eine vorrangige Bedeutung zu (Bybee, 1997; Gunn et al., 2008; Kolstø et al., 2006; Osborne, 2014; Rafolt et al., 2019; Yacoubian, 2015; Zeidler, 2014). Gleichzeitig sind SSI allgegenwärtig in der Lebensrealität zu finden (Zeidler & Kahn, 2014; Zeidler & Lewis, 2003), was neben den Vorteilen für den Erhebungsprozess auch die Alltagsnähe und damit Relevanz sowie die Angemessenheit des Erhebungsinstrumentes unterstreicht (Döring & Bortz, 2016; Lamnek & Krell, 2016).

Die Auswahl der Kontexte, die für die Verwendung in den Erhebungen oder Fördermöglichkeiten gewählt wurden, erfolgte entlang spezifischer Voraussetzungen, um den Anforderungen der Untersuchung gerecht zu werden. Zum einen sollten die Kontexte eine chemisch-fachwissenschaftliche Grundlage aufweisen, da die Anwendung naturwissenschaftlichen Wissens nach wie vor im Zentrum des Erkenntnisinteresses stand, zum anderen aber auch aktuell und politisch relevant sein, um einen konkreten Lebensweltbezug zu schaffen und die partizipatorischen Auswirkungen von naturwissenschaftsbezogenen Entscheidungen zu unterstreichen. Im gleichen Zug sollten den verwendeten Themen auch gesellschaftliche und ethische Implikationen innewohnen, um den kontroversen und uneindeutigen Charakter realer Fragestellungen widerzuspiegeln, die neben einer fachlichen auch eine moralisch-ethische Reflexion erfordern. Diese Vielschichtigkeit war für die Untersuchung der Multidimensionalität des Kritischen Denkens in derlei Bewertungsprozessen eine wichtige Voraussetzung, da neben der kognitiven auch die affektive Ebene konkreter Forschungsgegenstand war.

Insgesamt wurden im Rahmen der Untersuchungen neun verschiedene Kontexte eingesetzt, die erstmals in der Fragebogenerhebung S1.F1 der Studie S1 Anwendung fanden. Fünf dieser Kontexte (A bis E) dienten in den Interviews als Ausgangspunkt für die Problematisierung. Zusätzlich wurden diese fünf Kontexte gemeinsam mit drei weiteren SSI in der Systemisch-reflexiven Stellungnahme als wählbare Bezugsrahmen angeboten. Innerhalb der Erhebungen wurde sichergestellt, dass keine Settings im Interview gewählt wurden, welche die Teilnehmenden bereits im Rahmen der SrS in der Vermittlungseinheit bearbeiteten. Im Folgenden werden diese SSI-Themenbereiche kurz erläutert und auf deren naturwissenschaftliche sowie gesellschaftliche Bedeutung eingegangen.

4.2.6.1 Initialkontext: Handykauf

Den ersten eingesetzten Kontext bildete ein fiktiver Handyneukauf, bei welchem neben Neugeräten auch gebrauchte, erneuerte und überarbeitete Handys zur Verfügung stehen. Dieser Kontext eignet sich besonders für fachdidaktische Untersuchungen, da er den alltäglichen Konsum als Ausgangspunkt bietet, um komplexe Themen, wie Rohstoffverbrauch, Recycling und nachhaltigen Konsum, greifbar zu machen (Nordmann et al., 2015). Die Fragestellung wurde wie folgt formuliert:

> Ausgangssituation:
>
> Angenommen, Ihr aktuelles Handy wäre kaputt und nicht mehr zu retten. Sie stehen nun vor der Entscheidung, auf welche Weise Sie sich ein neues Gerät zulegen.
> Ihnen stehen verschiedene Kaufoptionen zur Verfügung, darunter auch gebrauchte oder überarbeitete Handys wie von Rebuy, Backmarket u. v. m..
>
> Aufgabe:
>
> Versuchen Sie, die Aspekte, die für Sie relevant sind, um eine Kaufentscheidung zu treffen, als Frage zu formulieren, die Sie sich selbst stellen.

Diese Situation stellt eine Alltagssituation im Leben der meisten Studierenden dar, da unsere heutige Technologisierung stetig fortschreitet. Damit einhergehend ist auch ein immer größer werdender Bedarf an relevanten Rohstoffen, wie den sogenannten Seltenen Erden, welche auch als kritische Rohstoffe bezeichnet werden. Die Kritikalität bezieht sich dabei vorrangig auf das Risiko von Versorgungsengpässen und weniger auf ihre tatsächliche Häufigkeit (Europäische Kommission, 2011). Unter den Seltenen Erden werden die Oxide der Seltenerdelemente verstanden, wozu

die Elemente Yttrium, Scandium, Lanthan und die darauf folgenden 14 Lanthanoide gehören. Diese Elemente werden auch als f-Metalle bezeichnet, da ihre atomare Struktur durch die schrittweise Besetzung der f-Orbitale charakterisiert ist (Reiners, 2003, 2017). Aufgrund ihrer optischen, magnetischen und katalytischen Eigenschaften sind sie besonders für die Hightechindustrie von großem Interesse und aus dem heutigen Alltag nicht mehr wegzudenken. Seltenerdelemente kommen häufig in Verbindung mit Radionukliden, wie Uran und Thorium, vor (Erdmann & Bundesanstalt für Geowissenschaften und Rohstoffe, 2021) und ihre Förderung ist mit erheblichen Umweltproblemen verbunden, da große Mengen an toxischen Chemikalien für die Extraktion und Trennung der Elemente benötigt werden. Bei der chemischen Extraktion aus Bastnasit oder thoriumhaltigen Monazit werden unter anderem Schwefelsäure, Salzsäure und direkte Chlorierung eingesetzt, wobei bis zu 50 % der Rohstoffe verloren gehen und als radioaktive Nebenprodukte anfallen. Diese chemischen Prozesse sind sehr energieaufwendig und führen zu großen, schwer zu entsorgenden Abfallmengen und damit einhergehend erheblichen Umweltschäden in den Abbaugebieten (Talens Peiró & Villalba Méndez, 2013). Hinzu kommt die politische Abhängigkeit von einigen wenigen Produzierenden. Über 80 % der abgebauten Seltenen Erden kommt aus den Bastnasitvorkommen Chinas, davon wiederum über 80 % aus der weltweit größten Lagerstätte in Bayan Obo in der inneren Mongolei (Deutscher Bundestag, 2022). Auch das Recycling von technischen Geräten und die Rückgewinnung der Seltenerdelemente gestaltet sich bisher als herausfordernd, da diese in den verwendeten Geräten mit anderen Materialien verbunden und schwer zu trennen sind (Marschall & Holdinghausen, 2018; Prechtl, 2022).

Die vermeintlich einfache Frage nach der Herkunft des neuen Mobilfunkgeräts wird damit zu einer mehrdimensionalen Herausforderung, bei welcher zwischen dem persönlichen Gefühl, welches mit einem Neugerät verbunden sein kann, und dem aktiven Beitrag zur Nachhaltigkeit abgewogen werden muss.

Auch in der Chemiedidaktik wird das Thema der Seltenen Erden mittlerweile verstärkt fokussiert und eine Sensibilisierung angestrebt (Kanbur et al., 2022; Lawatscheck et al., 2022; Mieger & Prechtl, 2017; Nordmann et al., 2015; Prechtl, 2017; Prechtl & Legscha, 2022; Reiners, 2017), wobei besonders das reflexive Potenzial einer Thematisierung der Seltenerdelemente hervorgehoben wird (Prechtl, 2022). Aus diesem Grund sowie der großen Alltagsnähe stellte die Kontextualisierung durch Seltenerdelemente einen geeigneten Ausgangspunkt für die Erprobung verschiedener Kontexte innerhalb der Erhebungen dar.

4.2.6.2 Interviewkontext A: Solarenergie vs. Windkraft

Das Setting der Gegenüberstellung von Solarenergie und Windkraft bietet zwei auf den ersten Blick nahezu gleichwertige erneuerbare Optionen der Energiegewinnung, deren Ausbau im Rahmen der Energiewende von der Bundesregierung gefördert wird (Bundesministerium für Wirtschaft und Klimaschutz (BMWK), 2024; Zahoransky, 2022b). Aus dieser Gleichwertigkeit ergibt sich die Notwendigkeit einer genaueren Betrachtung, denn trotz der zunächst guten Zukunftsaussichten beider Technologien sind diese nicht frei von Herausforderungen, welche die Umwelt und die Gesellschaft nachhaltig beeinflussen können.

Solarenergie wird hauptsächlich durch Photovoltaikanlagen gewonnen, bei denen Halbleitermaterialien, insbesondere aus monokristallinem oder polykristallinem Silizium, Lichtenergie durch den photovoltaischen Effekt direkt in elektrische Energie umwandeln (Bollin, 2022; Bründlinger et al., 2020; Jülch et al., 2015). Dieser beruht auf der Fähigkeit der Halbleitermaterialien, durch Licht-absorption Elektronen freizusetzen, die dann als Strom nutzbar gemacht werden können. Um die Effizienz der Solarzellen zu maximieren, werden oft spezielle Dotierstoffe, wie Phosphor, Gallium oder Bor, verwendet, welche die elektrische Leitfähigkeit des Siliziums erhöhen (Bollin, 2022; Bründlinger et al., 2020; Von Münch, 1993). Darüber hinaus werden Verbindungshalbleiter, wie Cadmiumtellurid (CdTe) oder Galliumarsenid (GaAs), eingesetzt, um die Effektivität weiter zu verbessern. Diese Materialien sind jedoch teilweise stark toxisch, was ihre Verwendung mit Blick auf die Umweltverträglichkeit und das Recycling problematisch macht (Bründlinger et al., 2020). Die Lebensdauer der Solarzellen beträgt zwischen 20 und 30 Jahre und aktuelle Recyclingverfahren sind bisher nur bedingt effizient, was sich in einer unvollständigen Rückgewinnung der eingesetzten Stoffe äußert.

Der größte Vorteil der Solarenergie liegt in ihrer Flexibilität und Dezentralisierung, da Photovoltaikanlagen auch auf Dächern und kleineren Flächen installiert werden können (Bollin, 2022). Damit tragen sie zur energetischen Unabhängigkeit auf privater sowie staatlicher Ebene bei und beinhalten das Potenzial zur Teilhabe und der Einbindung der*s Einzelnen in den Prozess der Energiewende. Gleichzeitig verursacht die Installation großflächiger Solarparks Konflikte, da sich diese auf die Biodiversität auswirken und zu Habitatveränderungen führen (Gasparatos et al., 2017). Auch ein Konflikt in Bezug auf die Nutzung von Agrarflächen und damit der Nahrungsmittelproduktion besteht (Pfeiffer et al., 2023).

Die ethisch-moralischen Herausforderungen bei der Verwendung der Solartechnik sind insbesondere in den bisher ineffizienten Recyclingprozessen begründet. Gerade beim Recycling und der Entsorgung von schwermetallhaltigen Solarmodulen können diese Stoffe in die Umwelt gelangen und diese nachhaltig schädigen (Bründlinger et al., 2020). Dadurch entsteht eine Diskrepanz zwischen dem Nutzen von Solarenergie als emissionsfreie Energiequelle und den ökologischen Risiken, welche mit der Entsorgung der Bestandteile einhergehen. Die unklare Abfallbewältigung und die fehlende Infrastruktur für ein umfassendes Recycling erscheinen damit als die größte Schwierigkeit bei der Verwendung von Solarzellen.

Die Gewinnung von Windenergie erfolgt hingegen durch moderne Windturbinen, die in ihrem Inneren Permanentmagnete enthalten. Diese enthalten große Mengen Seltenerdelemente, wie Neodym oder Dysprosium, deren Gewinnung mit den im Kontext Handykauf (vgl. Kap. 4.2.6.1) bereits erwähnten Herausforderungen einhergeht (Hartmann et al., 2015). Zusätzlich zum großen Bedarf an Seltenen Erden für die Permanentmagnete enthalten die Rotorblätter von Windkraftanlagen glasfaserverstärkte Kunststoffe, welche eine Lebenserwartung von 20 bis 25 Jahren haben und aufgrund der Materialkombinationen nur sehr aufwendig zu recyceln sind (Fichtner, 2022). Besonders Offshore-Anlagen stellen eine logistische Herausforderung dar, da deren Rückbau aufgrund der dezentralen Lage zusätzliche Kosten mit sich bringt (ebd.).

Obwohl Windkraftanlagen große Mengen Energie produzieren und damit nicht unerheblich zur Energiewende beisteuern (Bundesministerium für Wirtschaft und Klimaschutz (BMWK), 2024), leisten Anwohner der Installationsgebiete häufig Widerstand. Diese Konflikte entstehen durch die optischen Beeinträchtigungen der Landschaft, die Lärmbelästigung und die Auswirkungen auf die lokale Tierwelt, insbesondere Vögel, wie Rotmilane oder Auerhühner, und Fledermäuse (Brettschneider & Müller, 2020).

Die Abhängigkeit der Windkraft von Seltenen Erden und die Herausforderungen beim Rückbau und Recycling werfen ethische Fragen auf, nicht zuletzt, da die Rohstoffe unter kritischen Umwelt- und Arbeitsbedingungen gewonnen werden. Die Umweltzerstörung und gesundheitlichen Risiken für die lokale Bevölkerung sind erheblich und auch die damit einhergehende politische Abhängigkeit ist nicht unbedenklich. Gleichzeitig ist der technologische Fortschritt auf die Verwendung der Seltenen Erden angewiesen und besonders in Windkraftanlagen leisten diese einen großen Beitrag zur Energiewende.

Beide Energiegewinnungstechnologien, Solarenergie und Windkraft, zeichnen sich durch großes Potenzial bei gleichzeitig großen Risiken und Herausforderungen aus. Beide wirken auf den ersten Blick vielversprechend und umweltschonend, da die Auswirkungen produktionsbedingt vermeintlich unsichtbar ausgelagert sind. Dennoch bergen diese Technologien bisher ungelöste Probleme, denen ebenfalls begegnet werden muss. Beide Kontexte ähneln sich stark, unterscheiden sich jedoch hinsichtlich ihrer chemischen Implikationen und damit ihrer Auswirkungen. Aufgrund der mit einer Bewertung einhergehenden Herausforderung wurde daher folgende Problemstellung formuliert:

A1. Gehen wir von folgendem Setting aus: Du möchtest dich zu einem **Diskurs positionieren,** beispielsweise zur politischen Teilhabe im Rahmen einer Wahl. Es stehen dabei **zwei Positionen einander gegenüber**: Die eine Position steht für die Erweiterung der **alternativen Energieversorgung** durch **Windkraftanlagen**, die andere Position für den **Ausbau der Solarenergie**. Könntest du jetzt schon sagen, **zu welcher Position** du tendieren würdest?

Auf Grundlage dieser Problematisierung wurde in Studie S2 noch ein weiteres Setting formuliert, welches zusätzlich zu der Energiegewinnung aus Solarenergie und Windkraft die Optionen der Aufrechterhaltung der Atomenergie sowie die Aufrechterhaltung der Kohleverbrennung zur Wahl stellte. Ziel war es zu untersuchen, ob die gesellschaftlich stark emotional aufgeladenen Positionierungen gegenüber Atomenergie und Kohleverbrennung eine unbewusste Beeinflussung und damit eine stärkere Vorprägung der Antworten bewirken, die zu weniger reflektierten Entscheidungen führen könnten. Dieses Setting wurde aber nur im Rahmen der Vorstudie S2 erprobt und fand in späteren Studien keine Anwendung. Dennoch sollen die Rahmenbedingungen hier kurz umrissen werden.

A2. Gehen wir mal von folgendem Setting aus: Es stehen bald wieder Landtagswahlen an und **du möchtest deine Stimme abgeben**. Derzeit schwankst du zwischen mehreren Parteien. Die erste positioniert sich für die Erweiterung der alternativen Energieversorgung durch **Windkraftanlagen,** die zweite positioniert sich für den **Ausbau der Solarenergie**, die dritte positioniert sich für die **Aufrechterhaltung der Atomenergie** und die vierte für die Aufrechterhaltung der Kohleverbrennung. Könntest du jetzt schon sagen, **welcher Partei** du **deine Stimme** geben würdest und warum?

Eine genauere Betrachtung der Atomenergie ist bei der späteren Betrachtung der Atomenergie als Einzelkontext zu finden (vgl. Kap. 4.2.6.9).

Die Energiegewinnung aus Kohle basiert primär auf ihrer Verbrennung, welche chemisch exotherm abläuft. Der Hauptbestandteil von Kohle ist Kohlenstoff, der bei der Verbrennung mit Sauerstoff reagiert und dabei Kohlenstoffdioxid bildet. Mit der so gewonnenen Energie wird in Dampfkraftwerken Wasser erhitzt, um große Turbinen und daran gekoppelte Generatoren antreiben zu können (Zahoransky, 2022a).

Die Verbrennung von Kohle läuft jedoch nicht immer vollständig ab, was zur Bildung von Kohlenstoffmonoxid führt. Zusätzlich zur Verbrennung von Kohlenstoff enthält Kohle auch weitere Elemente, wie Schwefel und Stickstoff, die ebenfalls mit Sauerstoff reagieren und bei unvollständiger Verbrennung Verbindungen, wie Schwefeldioxid und Stickstoffoxide, bilden. Diese Abgase sind neben den CO_2-Emissionen die Hauptquelle von Luftverschmutzungen und saurem Regen (Schmidt et al., 2023), weshalb versucht wird, diese durch Entschwefelungs- und Entstickungsanlagen zu reduzieren (Zahoransky, 2022a).

In fortschrittlichen Kombinationskraftwerken, sogenannten Integrated Gasification Comnined Cycles (IGCC), wird die Kohle unter kontrollierter Sauerstoffzufuhr nicht vollständig verbrannt, sondern vergast (Markewitz et al., 2015). Dabei entsteht ein Synthesegas, welches hauptsächlich aus Kohlenstoffmonoxid und Wasserstoff besteht, aber auch geringe Mengen an Kohlenstoffdioxid, Schwefelwasserstoff und Stickstoff enthält. Dieses Synthesegas wird nach einer Aufreinigung in einer Gasturbine verbrannt und damit Generatoren angetrieben. Damit sind sie effizienter und umweltfreundlicher als herkömmliche Kohlekraftwerke, vor allem, wenn eine CO_2-Abscheidung integriert wird (Zahoransky, 2022a).

Dennoch bleiben die hohen CO_2-Emissionen eine große Belastung für die Umwelt und tragen maßgeblich zum Treibhauseffekt und damit zur globalen Erwärmung bei (Schmidt et al., 2023). Schwefeldioxid führt in der Umwelt zur Versauerung von Böden und Gewässern, was sich neben dem Einfluss auf das Ökosystem auch auf die Landwirtschaft auswirkt (Schmidt et al., 2023). Überdies schneidet die Kohleförderung selbst stark in die Landschaften ein, beeinträchtigt die Biodiversität und hat Boden- sowie Wasserverschmutzung zur Folge (Lohmann & Podbregar, 2012). Neben den Umweltrisiken hat der Abbau auch soziale Folgen, besonders dann, wenn die Kohle in Entwicklungsländern unter gefährlichen Arbeitsbedingungen abgebaut wird (ebd.).

Vor allem aus den ökologischen Gründen beschloss Deutschland schon 2020 die schrittweise Stilllegung aller Kohlekraftwerke bis 2038 und orientiert sich dabei am

Atomausstieg (Thrän et al., 2023). Dennoch stellt die Kohleverstromung global gesehen weiterhin eine relevante Energiequelle dar. Dies gilt insbesondere für Regionen mit hohem Energiebedarf, in denen umsetzungsfähige Konzepte für erneuerbare Energien und Speichertechnologien noch fehlen, wie etwa in vielen asiatischen Ländern. Hier dient Kohle nach wie vor als stabile Grundlastquelle (Markewitz et al., 2015).

Die schwerwiegenden Auswirkungen auf die Umwelt und die damit einhergehenden Auswirkungen auf folgende Generationen stellen die größten Gefahren dar, die den finanziellen Interessen und dem kurzfristigen Sicherstellen von Arbeitsplätzen und wirtschaftlicher Stabilität gegenüberstehen. Die Entscheidung, Kohle weiterhin als Energiequelle zu nutzen, stellt also ein Dilemma dar, bei dem kurzfristige finanzielle Vorteile gegen die langfristigen ökologischen und ethischen Konsequenzen abgewogen werden müssen.

Das Erweitern der Wahlmöglichkeiten um die Atom- und Kohleenergie lag in der Erprobung begründet, ob subjektiv klare Abgrenzungen in Entscheidungsprozessen zu einer direkteren und weniger überlegten Positionierung führen. Da die Atomenergie und die Kohleverbrennung gesellschaftlich stark negativ konnotiert und emotional geladen sind (Roeser, 2011), stellten diese beiden Optionen einen Kontrast zu den beiden erneuerbaren Energiequellen dar und erfüllten damit die Zielsetzung.

4.2.6.3 Interviewkontext B: E-Mobilität vs. Brennstoffzellenantrieb

Wie bei der Energiegewinnung gibt es auch im Hinblick auf den Straßenverkehr mehrere Ansätze zur langfristigen Reduktion von Emissionen. Die Gegenüberstellung von Elektromobilität und Wasserstoffbrennstoffzellentechnik umfasst dabei die beiden bekanntesten Konzepte, welche bereits im Alltag Verwendung finden.

Die Funktionsweise der Elektromobilität fußt auf den elektrochemischen Prozessen von Batterien, besonders von Lithium-Ionen-Batterien, die derzeit die am weitesten verbreitete Technologie darstellen (Karle, 2015; Röth et al., 2013; Sauer & Ringbeck, 2024). Es handelt sich dabei um sogenannte Interkalationsbatterien, was sich auf das Diffundieren des Lithiums in die bestehenden Kristallstrukturen von Kathode und Anode bezieht (Sauer & Ringbeck, 2024). Lithium zeichnet sich besonders durch sein leichtes Gewicht bei gleichzeitig hohem elektrochemischem Potenzial aus, was es für die Autoindustrie sehr attraktiv macht. Die Kathode der Batterien besteht häufig aus Metalloxiden, die aus verschiedenen Verhältnissen von Aluminium, Mangan, Kobalt und Nickel bestehen können, und häufig als NMC (Nickel, Mangan, Cobalt)

abgekürzt werden. Diese Kathodenverbindungen weisen eine hohe Energiedichte und Stabilität auf, beinhalten aber ökologisch und ethisch problematische Elemente, wie Kobalt, welches schwerpunktmäßig in Zentralafrika gewonnen wird (DECHEMA, 2015) und aufgrund seiner hohen Toxizität einen negativen Einfluss auf die Gesundheit der dort lebenden Bevölkerung hat (Banza et al., 2009). Als Anoden-Material wird meist Kohlenstoff in Form von Graphit in verschiedenen Modifikationen eingesetzt und in Teilen durch Silizium ergänzt. Dadurch erhöht sich die Energiedichte, weil sich mehr Lithium-Ionen in die stabilen Schicht-Strukturen einlagern können und die Batterie damit auch über viele Ladezyklen hinweg beständig machen (Sauer & Ringbeck, 2024). Während des Ladevorgangs werden die Lithium-Ionen aus der Kathode freigesetzt und bewegen sich durch den Elektrolyten zur Anode, wo sie in den Graphit-Schichten interkaliert werden. Beim Entladen wandern die Ionen wieder zurück zur Kathode und erzeugen dabei elektrische Energie, die dann in mechanische Energie zum Antrieb umgewandelt werden kann.

Neben den Lithium-Ionen-Batterien werden auch innovative Speicherkonzepte erforscht, mit dem Ziel, eine höhere Energiedichte und Lebensdauer zu ermöglichen. Lithium-Schwefel-Batterien beispielsweise stellen günstigere und leichtere Alternativen dar, weisen aber ein größeres Volumen und eine kürzere Zyklenlebensdauer auf, da die Kathodenmaterialien schneller degenerieren. Festkörperbatterien wiederum verwenden feste Elektrolyten und bieten Vorteile hinsichtlich Sicherheit und Energiedichte, sind aber in der Herstellung komplex und teuer. Eine vielversprechende Zukunftstechnologie der Speichermöglichkeiten ist die Lithium-Luft-Batterie, welche ein hohes Energiespeicherpotenzial und geringes Gewicht aufweist. Jedoch bildet sich in Teilen der Batterie Singulett-Sauerstoff, der die Kathode korrodiert. Die aktuelle Forschung sucht nach Möglichkeiten, die Bildung von Singulett-Sauerstoff zu verhindern oder andere Metalle als Lithium, unter anderem Natrium oder Kalium, einzusetzen (Sauer & Ringbeck, 2024).

Die großen Herausforderungen der Elektromobilität liegen vor allem im Recycling von Batterien. Da die Herstellung von Lithium-Ionen-Batterien energieintensiv ist und kritische Rohstoffe erfordert, ist die Rückgewinnung der Materialien, wie Lithium, Nickel oder Kobalt, ein entscheidender Faktor für die Nachhaltigkeit der Technologie. Die aktuellen Recyclingtechnologien bieten bisher nur unzureichende Wege zur Rückgewinnung, da verschiedene Kathodenmaterialien und Zusammensetzungen sowie das Zusammenwirken wasserfreier Elektrolyte den Recyclingprozess komplex erschweren (DECHEMA, 2015).

Die Elektromobilität gilt als Schlüsseltechnologie zur Reduktion von CO_2-Emissionen, jedoch argumentieren Kritiker, dass die CO_2-Emissionen vom Betrieb der Fahrzeuge nur zur Energiebereitstellung verlagert werden (Fläming et al., 2023). Damit ist nicht nur relevant, wie nachhaltig die Herstellung von Batterien und Elektrofahrzeugen möglich ist, sondern inwiefern der benötigte Strom aus erneuerbaren Energiequellen kommt. Dennoch führt die Einführung der Elektromobilität zu Verbesserungen der Luftqualität, insbesondere im städtischen Raum (Siebenpfeiffer, 2021). Auf Seite der Nutzenden gibt es neben den geringeren Betriebskosten im Vergleich zum Verbrenner aber auch Schattenseiten, wie die längeren Ladezeiten, die in erster Linie bei langen Reisen zu Verzögerungen führen können. Außerdem sind Batterien temperaturempfindlich, was das Laden bei niedrigen Temperaturen erschwert (Klell et al., 2018).

Eine Alternative zur Elektromobilität findet sich in der Verwendung von Wasserstoff als Energieträger. In den vergangenen Jahren wurde die Forschung und Entwicklung der Wasserstofftechnologie merklich verstärkt, um neben der Elektromobilität weitere nachhaltige Mobilitätslösungen zu fördern. Wasserstoffbetriebene Fahrzeuge stoßen lediglich Wasserdampf aus, was sie aus ökologischer Sicht zu einer ernst zu nehmenden Option macht (Mohrdiek et al., 2017). Der Antrieb wasserstoffbetriebener Fahrzeuge erfolgt über Wasserstoffbrennstoffzellen, in denen Wasserstoff an der Kathode mit Sauerstoff aus der Umgebungsluft reagiert. Auf Anodenseite wird Wasserstoff zur Anode der Brennstoffzelle geleitet und mittels Katalysatoren, zum Beispiel Platin, in Protonen (H^+) und Elektronen (e^-) aufgespalten. Die freigesetzten Elektronen fließen über einen externen Stromkreis zur Kathode und treiben durch die so entstehende elektrische Energie einen Elektromotor an. An der Kathodenseite reagieren die Protonen nach der Diffusion durch eine protonenleitende Membran, beispielsweise eine Polymerelektrolytmembran (PEM), mit dem in der Luft befindlichen Sauerstoff und den Elektronen. Diese Reaktion führt zur Bildung von Wasser, das als Wasserdampf aus der Brennstoffzelle abgeführt werden kann (Klell et al., 2018). Die Reaktion verläuft exotherm, wobei die Wärme entweder abgeführt oder zur Erhöhung der Systemeffizienz genutzt werden kann. Der große Vorteil der Wasserstofftechnik liegt in der hohen gravimetrischen Energiedichte des Wasserstoffs und der Trennung zwischen Energiewandler und -speicher, was die Reichweite der Fahrzeuge, verglichen mit batteriebetriebenen Autos, erhöht. Auch ist eine Betankung schneller möglich, was Wartezeiten an Tanksäulen reduziert (Klell et al., 2018). Der Wasserstoff muss aber zur Betankung komprimiert oder verflüssigt werden, wobei Drücke von bis zu 700 Bar oder Temperaturen von bis zu –253 °C

notwendig sind, die hohe Ansprüche an die verwendeten Materialien stellen (Mohrdiek et al., 2017; Schmidtchen & Wurster, 2017).

Wasserstoff bietet das Potenzial, als emissionsfreier Kraftstoff einen großen Beitrag zur Reduktion von Treibhausgasen zu leisten. Im Gegensatz zur batterieelektrischen Mobilität hat Wasserstoff aufgrund der geringen Betankungszeit und den höheren Reichweiten besonders im Hinblick auf Schwerlastverkehr und Langstreckenfahrzeuge klare Vorteile (Klell et al., 2018). Die Hürde einer vollständigen Umsetzung der Wasserstoffenergie ist jedoch groß, da der Aufbau einer komplett neuen Infrastruktur notwendig wäre und das hohe Kosten mit sich bringen würde. Auch ist die Gewinnung von Wasserstoff mit einem höheren Energiebedarf verbunden, welcher idealerweise selbst durch erneuerbare Energien gedeckt wird (grüner Wasserstoff). Aktuell stammt jedoch ein Großteil des Wasserstoffs aus der energieintensiven Dampfreformierung von Erdgas, was seinerseits wieder CO_2-Emissionen verursacht und die Umweltbilanz verschlechtert (Hamacher, 2017; Klell et al., 2018). Die Produktion von grünem Wasserstoff setzt eine vorhandene Energieinfrastruktur voraus, die fast vollständig auf erneuerbaren Energien basiert. Andernfalls könnte es zu einer Ressourcenkonkurrenz mit anderen Sektoren, wie dem Wohnungsbau oder industrieller Nutzung, kommen, was wiederum die Frage nach der Verteilung erneuerbaren Energien aufwirft. Dennoch wohnt dem Wasserstoff als Energiespeichermedium das Potenzial inne, als Puffer bei Stromspitzen grünen Wasserstoff herzustellen und die überschüssige Energie auf diese Weise zu speichern (Klell et al., 2018).

Insgesamt zeigt sich, dass sowohl batterieelektrischen als auch wasserstoffbetriebenen Antrieben ein großes Potenzial innewohnt, einen entscheidenden Beitrag zur Energiewende und Nachhaltigkeit leisten zu können. Beide Technologien weisen dabei beachtenswerte Vorteile, aber auch Nachteile auf, die eine genauere Betrachtung zur Abwägung und Positionierung erfordern. Das daraus resultierende Setting wurde wie folgt formuliert:

B. Gehen wir von folgendem Setting aus: Du möchtest dich zu einem **Diskurs positionieren**, beispielsweise zur politischen Teilhabe im Rahmen einer Wahl. Es stehen dabei **zwei Positionen einander gegenüber**: Die eine Position steht für die Subventionierung von **Elektrofahrzeugen**, die andere Position für die Subventionierung von **Brennstoffzellenfahrzeugen**. Könntest du jetzt schon sagen, **zu welcher Position** du tendieren würdest?

Damit stellt dieses Setting eine aktuelle, politisch sowie gesellschaftlich relevante Problemstellung dar, der sich langfristig im Rahmen der gesellschaftlichen Partizipation gestellt werden muss. Nur durch eine multiperspektivische Betrachtungsweise kann eine umfassende Bewertung dieses Kontextes stattfinden und eine begründete Positionierung erfolgen.

Dieses Setting wurde in den Interviewerhebungen der Studie S3, in der SrS der Vermittlungseinheit der Studie S5 sowie in der Fragebogenerhebungen der Studie S5 verwendet.

4.2.6.4 Interviewkontext C: Plastik

Plastik ist ein Sammelbegriff für synthetische Werkstoffpolymere, die je nach Zusammensetzung verschiedene mechanische und chemische Eigenschaften aufweisen können (Hengstmann & Tamminga, 2022; Laskar & Kumar, 2019). Sie bestehen aus Monomeren, hauptsächlich kohlenstoffbasierten Einheiten, welche über eine Polymerisation auf verschiedene Weise chemisch miteinander verknüpft werden können und in verschiedenen Verfahren, wie Extrusion oder Spritzguss, verarbeitet werden. Die dafür benötigten organischen Stoffe entspringen meist fossilen Quellen, wie zum Beispiel Erdöl, und werden aus nicht erneuerbaren Rohstoffen hergestellt (Eyrer, 2020; Hengstmann & Tamminga, 2022; Martens & Goldmann, 2016). Zusätzlich enthalten Kunststoffe häufig Additive, um bestimmte Eigenschaften, wie Farbe, Härte oder Brennbarkeit, zu beeinflussen. Diese Additive können jedoch durch Verwitterung oder Gebrauch aus dem Material freigesetzt und in die Umwelt abgegeben werden, was aufgrund ihrer Toxizität Gefahren für die Umwelt birgt (Eyrer, 2020; Fath, 2019; Hengstmann & Tamminga, 2022).

Kunststoffe lassen sich in zwei Hauptkategorien einteilen. Die Thermoplaste sind durch Erwärmen verformbar, was ein mehrfaches Schmelzen und Recyclen ermöglicht. Beispiele für Thermoplasten sind Polyethylen (PE), Polypropylen (PP) und Polyethylenterephthalat (PET). Die Duroplaste hingegen sind stark vernetzt und daher nach der Aushärtung nicht mehr formbar, was sie sehr widerstandsfähig macht (Eyerer, 2020; Goodship, 2007; Hopewell et al., 2009; Shen & Worrell, 2024).

Eine besondere Gefahr geht vom sogenannten Einmalplastik oder Einwegkunststoffen, wie Polyethylen (PE) oder Polypropylen (PP), aus, da aus ihnen Wegwerfverpackungen und Einmalprodukte hergestellt werden und sie gleichzeitig nur schwer biologisch abbaubar sind (Hengstmann & Tamminga, 2022). In der Umwelt zersetzen sie sich in sogenanntes Mikroplastik mit Partikelgrößen von 100 nm bis 5 mm (Fath,

2019; Napper et al., 2015). Diese Partikel fungieren häufig als Träger für Schadstoffe, wie Pestizide und giftige Chemikalien, die von den Plastikoberflächen adsorbiert werden und von Organismen, einschließlich des Menschen, aufgenommen und sogar im Gewebe gespeichert werden (Fath, 2019; Hengstmann & Tamminga, 2022; Laskar & Kumar, 2019). Das Recycling der Plastikabfälle stellt zwar eine potenzielle Lösung dar, ist aber mit großen technologischen Herausforderungen verbunden. Mechanische und thermische Einflüsse können die Polymere zerstören, was das Recycling auf werkstofflicher Ebene erschwert (Martens & Goldmann, 2016). Auch wenn die mechanische Zerkleinerung von Plastik die gebräuchlichste Methode bleibt, stellt die hohe Diversität von Kunststoffen und die Selektion im Recyclingprozess die größte Herausforderung dar. Die Verwendung von Additiven erschwert zudem den Zerkleinerungsprozess und die Gewinnung von Reinstoffen (Shen & Worrell, 2024). Das chemische Recycling ermöglicht es, Polymere durch Depolymerisation wieder in ihre Monomere aufzubrechen, benötigt aber große Mengen Energie (Hann & Connock, 2020; Shen & Worrell, 2024), was wiederum die Umweltbilanz verschlechtert. Zwar gibt es vermeintliche ökologische Alternativen, aber auch Bio-Plastik ist nicht vollständig biologisch abbaubar (Hengstmann & Tamminga, 2022).

Ein Verbot von Einwegplastik könnte das Konsumverhalten nachhaltiger gestalten und die Nachfrage nach Alternativen anregen (Herberz et al., 2020). Andererseits können Konsumierende das Verbot als Einschnitt in die individuelle Freiheit verstehen, was wiederum kontraproduktiv für die Sensibilisierung des eigentlichen Problems sein könnte (Bharadwaj et al., 2020). Gleichzeitig geht ein Einmalplastikverbot mit wirtschaftlichen Konsequenzen einher, bietet aber auch Chancen für neue Arbeitsfelder, wie die Entwicklung nachhaltiger Alternativen. Die Subventionierung des Recyclings wiederum könnte die Abfallinfrastruktur stärken, während gleichzeitig der Plastikverbrauch gesenkt werden müsste, um eine wirksame Alternative zu bieten. Die Entscheidung, Plastik zu verbieten oder das Recyclingsystem zu stärken, stellt sich als multiperspektivisches Problem dar, für das es keine eindeutige Lösung gibt (Godfrey, 2019). Während die Auswirkungen auf die Umwelt und die potenzielle Anreicherung von Schadstoffen in der Nahrungskette schwer wiegen, stehen diesen wirtschaftliche Herausforderungen und Einschnitte in das Konsumverhalten der Menschen gegenüber. Aufgrund dieser Vielschichtigkeit eignet sich dieser Kontext für eine Problematisierung besonders und fordert eine mehrdimensionale Betrachtung ein. Die Fragestellung des Settings wurde daher wie folgt formuliert:

C. Gehen wir von folgendem Setting aus: Du möchtest dich zu einem **Diskurs positionieren**, beispielsweise zur politischen Teilhabe im Rahmen einer Wahl.

> Es stehen dabei **zwei Positionen einander gegenüber**: Die eine Position steht für die stärkere Regulierung und Verbote von Einwegplastik, die andere Position stellt sich **gegen ein Verbot**, sondern spricht sich **für eine Subventionierung des Recyclings** aus. Könntest du jetzt schon sagen, **zu welcher Position** du tendieren würdest?

Dieser Kontext kam in der SrS der Vermittlungseinheit der Studie S3 sowie in der Interviewerhebung S5 zur Anwendung.

4.2.6.5 Interviewkontext D: Fracking

Fracking, auch Hydraulic Fracturing genannt, bezeichnet ein Verfahren zur Förderung von in tiefen Gesteinsformationen eingeschlossenem Öl und Erdgas, das seit 1947 angewandt und vor allem durch seinen massiven Einsatz in den USA bekannt wurde (Habrich-Böcker et al., 2015). Es ermöglicht den Zugang zu unkonventionellen Lagerstätten, wie Schiefergas, Kohleflözengas oder Tight Gas, die aufgrund der geringen Durchlässigkeit des Gesteins mit herkömmlichen Methoden nicht erreichbar wären. Um diese zu erschließen, werden beim Fracking durch das Hineinpressen von Frack-Flüssigkeiten Risse im Gestein erzeugt. Diese Frack-Fluide bestehen zu 70 bis 90 % aus Wasser, enthalten aber auch zugesetzte Stützmittel als Additive, um ihre Eigenschaften, wie die Viskosität oder das Wachsen von Bakterien, zu regulieren (von Tschirnhaus, 2022).

Trotz starker Sicherheitsregularien bleibt das Verfahren riskant. So enthält der Flowback, die nach dem Frackingprozess zurückgeführte Flüssigkeit, häufig eine Vielzahl geogener Stoffe, die aus den Tiefen des Gesteins hervorgebracht werden. Dazu zählen insbesondere Metallsalze, die in das Wasser übergehen können und problematisch für Umwelt und Wasserqualität sind, radioaktive Stoffe, wie Radium, oder auch organische Verbindungen, wie das toxische und krebserregende Benzol (Frimmel et al., 2012; Mortimer & Müller, 2010). Die Zusammensetzung des Flowbacks stellt damit eine erhebliche Herausforderung für die Entsorgung und die Sicherstellung der Umweltverträglichkeit dar und die Kontamination des Grundwassers ist eine mögliche Gefahr mit direkten Auswirkungen auf die umliegende Bevölkerung (Habrich-Böcker et al., 2015). Auch das Risiko anthropogen induzierter seismischer Erschütterungen ist gegeben (Schultz et al., 2020) und die Luft wird durch zusätzliche Methanemissionen oder smogbildenden Stickoxide belastet (Habrich-Böcker et al., 2015). Trotzdem gilt Fracking in Zeiten politischer Unruhen und der damit einhergehenden Unsicherheiten in der Rohstoffversorgung als Chance, die Energieversorgung zu sichern und die damit verbundene Wirtschaft zu stabilisieren (Habrich-

Böcker et al., 2015; von Tschirnhaus, 2022). Auch in Deutschland ist die Gewinnung von Schiefergas durch Fracking schon lange in der Diskussion, um die Importabhängigkeit zu reduzieren, wurde aber bereits 2014 aufgrund der Risiken für das Grundwasser und die Klimabilanz abgewiesen (Bertram et al., 2014). Trotz des Ukraine-Konflikts lehnt das Bundesamt auch heute noch den Einsatz von Fracking ab, da die Ressourcen nicht rechtzeitig verfügbar wären, um die kurzfristigen Versorgungsengpässe zu überbrücken, und setzt in seiner Energiestrategie weiterhin auf die Förderung von erneuerbaren Energien und grünem Wasserstoff (Bertram et al., 2022). Trotzdem bleibt die Forderung nach einer Neubewertung dieser Einschätzung bestehen, wobei besonders die wirtschaftlichen Vorteile hervorgehoben werden (von Tschirnhaus, 2022). Unter ethisch-moralischen Gesichtspunkten stellt sich die Frage, ob die wirtschaftlichen und politischen Vorteile die Risiken für die Umwelt und menschliche Gesundheit aufwiegen oder ob aufgrund der potenziellen Gefahren von dieser Förderpraxis Abstand genommen werden sollte.

Das Thema Fracking fand zudem bereits chemiedidaktische Beachtung und wurde als gesellschaftskritisch-problemorientierter Chemieunterricht in Unterrichtskonzepte in Deutschland und den USA übertragen (Zowada et al., 2018; Zowada & Eilks, 2018).

Die zu diesem Kontext zugehörige Fragestellung wurde wie folgt formuliert:

D. Gehen wir von folgendem Setting aus: Du möchtest dich zu einem **Diskurs positionieren,** beispielsweise zur politischen Teilhabe im Rahmen einer Wahl. Es stehen dabei zwei Positionen einander gegenüber: Die eine Position steht für die **Förderung von Erdgas** mittels **hydraulischer Frakturierung**, die andere Position spricht sich **gegen diese Art der Erdgasförderung** aus. Könntest du jetzt schon sagen, **zu welcher Position** du tendieren würdest?

Dieser Kontext wurde in der SrS der Studie S3 sowie in den Interviews der Studie S5 verwendet.

4.2.6.6 Interviewkontext E: Stickoxidgrenzwerte

Stickoxide (NO_X), wie beispielsweise Stickstoffdioxid, entstehen durch die Oxidation des in der Luft befindlichen Stickstoffs bei Verbrennungsprozessen. Sie sind bedeutende Schadstoffe vor allem in städtischen Gebieten, da sie von Fahrzeugmotoren, insbesondere Dieselmotoren, und der Industrie emittiert werden (Borgeest, 2017). Stickoxide wirken sich negativ auf die menschliche Gesundheit aus und können

Atemwegs-, Herz-Kreislauf- und Krebserkrankungen sowie eine erhöhte Sterblichkeit begünstigen (Heinrich, 2018; Helmers, 2009). Stickstoffdioxid wirkt sich überdies auf die Umwelt in Form der Förderung von saurem Regen aus und katalysiert im Sommer die Bildung von Ozon (Borgeest, 2017).

Zur Regulation des Schadstoffausstoßes wurden Grenzwerte eingeführt, welche eine Balance zwischen Gesundheitsanforderungen und den wirtschaftlichen sowie infrastrukturellen Faktoren schaffen sollen. Der zulässige Grenzwert der EU für Stickstoffdioxid liegt bei einem Jahresmittelwert von 40 $\mu g/m^3$ (Rat der Europäischen Union, 2008), welcher aber an vielen Messstandorten in Deutschland überschritten wird (Becker et al., 2009; Butler et al., 2018). Dabei liegt dieser Grenzwert sogar um ein Vielfaches höher als die Empfehlung der WHO von 10 $\mu g/m^3$ (Köhler et al., 2019; World Health Organization, 2021). Abgasreinigungsverfahren, wie die selektive katalytische Reduktion (SCR), unterstützen zwar die Reduktion der Emissionen, reichen jedoch vor allem in stark belasteten Gebieten nicht aus (Beckmann, 2011).

Dennoch gibt es auch Kritik an der aktuellen Datenlage und Zweifel an der Gefährlichkeit von Grenzwertüberschreitungen. Dabei werden die Aussagekraft und Güte epidemiologischer Studien kritisiert und die gesundheitlichen Risiken insbesondere für gesunde Personen infrage gestellt (Köhler et al., 2019).

Eine weitere Senkung der Grenzwerte könnte zu einer besseren Luftqualität und einer Reduktion der gesundheitlichen Risiken beitragen, würde aber zugleich einen erheblichen wirtschaftlichen Einschnitt bedeuten, da große Investitionen in neue Fahrzeugtechnologien und Mobilitätsstrukturen geleistet werden müssten. Die Beendigung der Dieselsubventionierung bietet eine Möglichkeit zur Schadstoffreduktion (Butler et al., 2018), würde aber die Fahrzeugindustrie stark belasten, da eine vollständige Neuausrichtung innerhalb kurzer Zeit nötig wäre. Die Elektromobilität stellt zwar bereits eine Alternative dar, ist aber bislang nicht für jeden erschwinglich und die Infrastruktur auf einen kompletten Umschwung nicht vorbereitet.

Im Gegensatz dazu wäre ein Beibehalten des Grenzwertes entgegen den Empfehlungen der WHO ein Risiko für die menschliche Gesundheit, was sich sowohl in der Lebensqualität als auch in den Gesundheitskosten niederschlagen würde. Aus diesem Dilemma heraus wurde das Setting wie folgt formuliert:

E. Gehen wir von folgendem Setting aus: Du möchtest dich zu einem **Diskurs positionieren**, beispielsweise zur politischen Teilhabe im Rahmen einer Wahl. Es steht dabei die Frage im Raum, **politisch die Grenzwerte für Stickoxide**

weiter herabzusetzen, was mit **Maßnahmen zur Stickoxidreduktion** zur Grenzwerteinhaltung einhergehen würde. Könntest du jetzt schon sagen, **ob du dich für oder gegen ein weiteres Herabsetzen der Stickoxidgrenzwerte positionieren würdest**?

Dieser Kontext wurde ausschließlich in der Interviewerhebung der Studie S5 verwendet.

4.2.6.7 Zusatzkontext 1: Nanotechnologie

Die Nanotechnologie macht sich die Eigenschaften von Materialien auf der Nanometerskala (1 – 100 nm) zunutze, da Materialeigenschaften in diesen Kleinstbereichen von den Gesetzen der klassischen Physik abweichen. Diese besonderen optischen, magnetischen und mechanischen Eigenschaften und eine erhöhte Reaktivität liegen maßgeblich im Verhältnis zwischen Oberfläche und Volumen begründet (Fahrner & Ulyashin, 2017) und sind besonders für die Industrie, Umwelttechnik und Medizin von großem Interesse. Ein Beispiel ist die gezielte Medikamentenverabreichung durch Nanomaterialien, die so modifiziert werden, dass sie biologische Barrieren überwinden können und Medikamente in spezifisches erkranktes Gewebe applizieren. Dieser Ansatz reduziert Nebenwirkungen und stellt einen großen Fortschritt auf dem Gebiet der Krebstherapie dar (McNeil, 2011). In der Umwelttechnik zeigen Nanopartikel zudem das Potenzial, durch ihre hohe Reaktivität Schadstoffe zu neutralisieren, indem sie als Katalysatoren für die Zersetzung toxischer Substanzen fungieren. So können nanoskalige Eisenpartikel effektiv zur Reinigung von Grundwasser und Sanierung kontaminierter Böden eingesetzt werden (Karn et al., 2009; Taran et al., 2020). Auch eröffnen Nanomaterialien neue Möglichkeiten im Bereich der erneuerbaren Energien, indem sie zur Verbesserung von Solarzellen oder Batterien beitragen (Hess & Lamprou, 2012).

Trotz ihrer vielen Anwendungsmöglichkeiten bergen Nanopartikel aufgrund ihrer Größe und Eigenschaften auch potenzielle Umweltrisiken, da sie sich besonders in freier Form in der Luft, dem Wasser und Boden verteilen und dadurch zu unvorhergesehenen ökologischen und gesundheitlichen Belastungen führen können. Sie dringen beispielsweise durch Inhalation in Organismen ein und können aufgrund ihrer Größe sogar die Blut-Hirn-Schranke überwinden, was zur Schädigung von Geweben und zur Entstehung von Tumoren führen kann (Hannah & Thompson, 2008). Die Ungewissheit über die potenziellen Langzeitrisiken und die geringe Finanzierung für

die Erforschung der Sicherheits- und Umweltaspekte, welche nur etwa 1 % der Investitionen in die Nanotechnologie ausmachen, verstärken die Unsicherheit und vermindern die gesellschaftliche Akzeptanz (ebd.).

Die Vorteile der Nanotechnologie sind beträchtlich und berühren viele Anwendungsfelder. Diese breite Anwendbarkeit beinhaltet das Vermögen zur Schaffung neuer Industrien und Arbeitsplätze, was die Nanotechnologie auch wirtschaftlich sehr interessant macht (Howard & Wetter, 2012). Gleichzeitig betreffen die Sicherheitsaspekte sowohl die Umwelt als auch die Gesellschaft und es kann die Frage gestellt werden, inwieweit eine präventive Regulierung notwendig ist, um etwaige Langzeitrisiken zu vermeiden (Besley et al., 2008). Damit wird auch das Vertrauen der Öffentlichkeit in die Wissenschaft beeinflusst, denn ein Mangel an Transparenz und der Eindruck unzureichender Kontrolle kann das Vertrauen in die Technologie selbst untergraben (Cobb & Macoubrie, 2004).

Es zeigt sich, dass Nanomaterialien große Möglichkeiten und ein großes Risiko in sich vereinen und eine Positionierung in diesem Kontext einer umfassenden Reflexion bedarf. Aus diesem Grund eignet sich die Nanotechnologie als problematisierender Aufriss und zur Förderung metakognitiver Denkweisen. Die verwendete Fragestellung lautete:

> Sollte die **Entwicklung und Anwendung von Nanotechnologie strenger reguliert** werden, um **potenzielle Risiken** für die Gesundheit und Umwelt zu **minimieren**, oder sollten wir die **Chancen und Potenziale der Nanotechnologie nutzen** und die Regulierung flexibel gestalten, um **Innovationen und wirtschaftlichen Fortschritt** zu ermöglichen?

Dieser Kontext wurde in beiden SrS der Vermittlungseinheiten der Studien S3 und S5 eingesetzt.

4.2.6.8 Zusatzkontext 2: Chemie in der Landwirtschaft

In der modernen Landwirtschaft spielen chemische Stoffe, wie Stickstoffdünger und Pflanzenschutzmittel, eine wichtige Rolle. Besonders Stickstoff ist bei der Synthese von Proteinen und für die Photosynthese von großer Relevanz und steigert so bei Substitution die Biomasseproduktion erheblich (Li & Gao, 2024). Jedoch werden Düngemittel häufig im Übermaß angewandt, das über den tatsächlichen Bedarf der Pflanzen hinausgeht. Dadurch gelangt überschüssiger Stickstoff in den Boden und das Grundwasser und kann durch Auswaschung in aquatische Ökosysteme fließen,

was die Eutrophierung der Gewässer und die Bedrohung der Biodiversität zur Folge haben kann (Haller et al., 2020; Reganold & Wachter, 2016). Gelangen die Nitrate ins Grundwasser, gehen damit neben den ökologischen Problemen direkte gesundheitliche Risiken für den Menschen einher (Hopp, 2018). Um dem entgegenzuwirken, wird bei der Düngemittelanwendung auch von der 4R-Strategie gesprochen: richtige Menge, richtige Zeit, richtiger Ort und richtige Quelle (Fixen & Johnston, 2012).

Neben den Düngemitteln ist auch der Einsatz von Pflanzenschutzmitteln, wie Glyphosat, in der Landwirtschaft weitverbreitet, um Schädlingsbefall zu minimieren. Dabei wirken diese Stoffe nicht selektiv auf einzelne Organismen, sondern beeinträchtigen auch andere Pflanzen und Tiere in ihrem Ökosystem, zum Beispiel Amphibien (R. M. Mann et al., 2009).

Das Sicherstellen der Welternährung bildet das zentrale Ziel der Landwirtschaft, auf welcher wegen der stetig wachsenden Bevölkerung und der Lebensmittelstandards ein immer größerer Druck lastet (Dachler, 2023). Chemische Düngemittel und Pflanzenschutzmittel sind in diesem Kontext nahezu unentbehrlich, da nur so das Pflanzenwachstum und die landwirtschaftliche Produktivität optimiert werden können. Es wird angenommen, dass die weltweite landwirtschaftliche Produktion bis 2050 um bis zu 70 % gesteigert werden muss und dieser Anteil in Entwicklungsländern noch höher liegt (ebd.). Die daraus resultierende Abhängigkeit von Dünge- und Pflanzenschutzmitteln sorgt für einen Konflikt mit dem Ziel des nachhaltigeren Anbaus aus ökologischer Landwirtschaft, welcher häufig ertragsschwächer ist (Haller et al., 2020). Das führt in Konsequenz zu einem Dilemma zwischen dem Bedarf an hohen Erträgen und den Folgen der verstärkten Nutzung von Pestiziden und Düngemitteln für die Umwelt, aber auch durch Nahrung und Wasser für die Gesundheit der Menschen (Frische et al., 2016).

Dieses Dilemma wurde für die Kontextualisierung als folgende Fragestellung formuliert:

> Sollten **chemische Pestizide und Düngemittel in der Landwirtschaft weiterhin weitverbreitet eingesetzt werden**, um Ernteerträge zu maximieren und die **Welternährungssicherheit** zu gewährleisten, oder sollten wir **verstärkt auf alternative Methoden**, wie biologischen Anbau und nachhaltige Landwirtschaft**, setzen**, um Umweltauswirkungen und **potenzielle Gesundheitsrisiken** zu minimieren?

Dieser Kontext wurde in der SrS in den Vermittlungseinheiten von Studie S3 und S5 verwendet.

4.2.6.9 Zusatzkontext 3: Atomenergie

Die Atomenergie spielt eine zentrale Rolle in der globalen Energieversorgung, wobei sich die gesellschaftliche Akzeptanz weltweit unterscheidet, nicht zuletzt aufgrund der Vorfälle in Fukushima und Tschernobyl (Roeser, 2011). In Deutschland hat die gesellschaftliche Abneigung bereits zum Atomausstieg geführt, wohingegen andere Nationen, wie Frankreich und Japan, die Technologie als unverzichtbare Quelle zur CO_2-Reduktion sehen (Nathaniel et al., 2021; Schnabel, 2020; Siqueira et al., 2019).

Die Energieproduktion in Kernkraftwerken basiert auf den kontrollierten Kernspaltungen des Uranisotops ^{235}U und des Plutoniumisotops ^{239}Pu, welche große Mengen an Energie freisetzen. In typischen Leichtwasserreaktoren wird spaltbares ^{235}U neutroneninduziert gespalten und die frei werdende Energie zur Erhitzung von Wasser verwendet, was durch den entstehenden Dampf in einer Turbine Strom erzeugt (Allelein, 2022; Kessler et al., 2014). Dabei entstehen viele radioaktive Spaltprodukte, wie Iod ^{131}I oder Cäsium ^{137}Cs, die eine lange Halbwertszeit und eine hohe Umwelttoxizität aufweisen. Die Entsorgung von Spaltprodukten und abgebrannten Brennelementen erweist sich als besonders schwierig, da sie eine Lagerung über einen Zeitraum von mehreren Jahrhunderten erfordern. Auch die Anreicherung des Urans bedarf komplexer chemischer Prozesse, wie die Überführung des Uranerzes Triuranoctoxid U_3O_8 in Uranhexafluorid UF_6 (ebd.), was ebenfalls erhebliche Umweltbelastungen mit sich bringen kann. Eine weitere Herausforderung ist die Nachzerfallswärme, die auch eine Kühlung nach Abschaltung des Reaktors nötig machen, um eine Kernschmelze zu verhindern (Kessler et al., 2014).

Die Uranversorgung birgt zusätzlich ein ethisches Problem, da der Abbau häufig in politisch instabilen Regionen stattfindet und zu Umweltschäden sowie sozialen Ungleichheiten führt (Bund für Umwelt und Naturschutz Deutschland et al., 2022).

Die Kernenergie bietet den Vorteil, große Energiemengen nahezu CO_2-frei zu erzeugen, was sie in vielen Ländern als wertvolle Übergangslösung in der Energiewende macht (Siqueira et al., 2019). Allerdings stellen die hohen Bau- und Rückbaukosten sowie die langfristige Endlagerung der radioaktiven Abfälle durchaus ernstzunehmende Probleme dar (Schnabel, 2020). Die ethische Bewertung der Kernenergie ist aufgrund ihrer potenziellen Risiken und Vorteile komplex und dabei hoch-

emotional geladen (Roeser, 2011). Das zentrale Dilemma betrifft vor allem die Perspektive auf zukünftige Generationen, da zugunsten der heutigen CO_2-Reduktion die Verantwortung für den Umgang mit den endgelagerten radioaktiven Abfällen an spätere Generationen übergeben wird (ebd.). Daraus ergab sich für die Kontexte folgende Fragestellung:

> Sollte **Atomenergie** als eine bedeutende Energiequelle in der Zukunft **weiter ausgebaut** werden, um den **steigenden Energiebedarf zu decken** und den Klimawandel zu bekämpfen, oder sollten wir **verstärkt auf erneuerbare Energien setzen** und Atomenergie aufgrund ihrer **potenziellen Risiken und Auswirkungen** einschränken oder sogar auslaufen lassen?

Dieser Kontext wurde in den Interviewerhebungen als Teil des Kontextes A2 sowie in der SrS der Vermittlungseinheiten der Studien S3 und S5 verwendet.

4.3 Methoden der Datenauswertung

Zur Sicherung der Transparenz und Förderung der Intersubjektivität werden in diesem Kapitel die Schritte der Datenverarbeitung offengelegt (Früh, 2017). Dies umfasst sowohl die Beschreibung der qualitativen Methoden als auch eine detaillierte Darstellung der quantitativen Auswertungen, um die Nachvollziehbarkeit zu gewährleisten und die Argumentationsstruktur klar darzulegen.

4.3.1 Qualitative Auswertungsmethoden

Die Interviews wurden nach der Durchführung und Aufnahme mithilfe des KI-Tools von *f4x* vortranskribiert. Die damit verbundene Verarbeitung der hochzuladenden Daten findet dabei DSGVO-konform und verschlüsselt statt (Dr. Dresing & Pehl GmbH, 2023). Nach der Vortranskription wurden die Daten in die Transkriptionssoftware *f4* übertragen (Dr. Dresing & Pehl GmbH, 2024; Dresing & Pehl, 2018; Gläser-Zikuda et al., 2022; Kuckartz, 2016), wo eine händische Überprüfung und Nachtranskription vorgenommen wurde. Die Überführung von Ton in Schrift fand zur Wahrung der Güte und Nachvollziehbarkeit regelgeleitet statt, um eine konsistente Grundlage für die spätere Inhaltsanalyse zu schaffen (Przyborski & Wohlrab-Sahr, 2021). Da vornehmlich der Inhalt des Gesagten und nicht das Verhalten oder die Art und Weise des Gesprochenen im Fokus stand, wurde eine modifizierte wörtliche Transkription in Anlehnung an Dresing und Pehl (2018, S. 22 f.) und Kuckartz (2016, S. 167) als ausreichend erachtet. Damit stellt die Transkription eine Reduktion

dar, welche weder Gestik und Mimik noch Intonation adäquat wiedergeben kann, sondern einen bewussten Fokus setzt (Przyborski & Wohlrab-Sahr, 2021). Die verwendeten Transkriptionsregeln sind im Anhang (A.2.1) einsehbar. Jedes Transkript durchlief nach der KI-Transkription und der ersten Nachtranskription zwei Prüfdurchgänge zur Korrektur und Wahrung der qualitativen Ansprüche (Dresing & Pehl, 2018). Die Speicherung fand zu Dokumentationszwecken sowohl im RTF-Format als auch final als PDF statt.

Zur Analyse der Interviewtranskripte standen verschiedene Methoden zur Verfügung, wobei grundsätzlich in rekonstruktive und strukturierende Verfahren unterschieden werden kann (Przyborski & Wohlrab-Sahr, 2021). Rekonstruktive Verfahren, wie die *Objektive Hermeneutik* oder *Grounded Theory,* versuchen tiefere Bedeutungsstrukturen oder implizite Sinninhalte von Texten, Aussagen oder Handlungen zu erschließen, mit dem Ziel, subjektive oder soziale Phänomene in ihrem Entstehungszusammenhang nachzuvollziehen (Böhm, 2019). Da diese Art der Rekonstruktion den Fokus auf das Verstehen und Interpretieren legt, sind damit verbundene Fragestellungen auf die Bedeutung von Daten ausgerichtet. Strukturierende Ansätze, wie die *qualitative Inhaltsanalyse,* hingegen legen den Schwerpunkt auf das Ordnen und Systematisieren der Daten zur Beantwortung von Fragestellungen, die sich am Inhalt und dessen Kategorisierung orientieren. Da in diesem Forschungsprojekt vorrangig Denkperspektiven untersucht werden sollen, liegt das Zentrum des Interesses bei der Ordnung und Systematisierung der gewonnenen Daten und damit bei der Kategorisierung und Perspektivierung, um die Strukturen des Denkens aufzeigen zu können. Weniger wird eine tiefergehende Analyse der Antworten mit der Zielsetzung der Ergründung der dahinterliegenden Sinnstrukturen angestrebt. Dennoch soll auch ein kurzer Überblick über die zwei bekanntesten rekonstruktiven Verfahren gegeben werden, um die Entscheidung für ein strukturierendes Vorgehen nachvollziehbar zu machen.

Die Objektive Hermeneutik konzentriert sich auf die Rekonstruktion latenter Bedeutungen und sozialer Sinnstrukturen durch die sequenzielle Analyse von Texten (Lamnek & Krell, 2016; Reichertz, 2016). Sie stellt eine detaillierte Betrachtungsweise von Textstellen dar, ist daher jedoch für die Analyse großer Datenmengen, wie bei Interviews, wenig praktikabel (Lamnek & Krell, 2016). Gleichzeitig ist die Objektive Hermeneutik stark interpretativ und beinhaltet bei den hier verfolgten Fragestellungen keinen zusätzlichen Erkenntnisgewinn, weshalb diese Methodik im Hinblick auf die Zielsetzung und die Praktikabilität für die Analyse der im Rahmen des Forschungsprojekts gewonnenen Daten als nicht zielführend angesehen wurde.

Die Methodik der *Grounded Theory* hingegen hat schwerpunktmäßig die Entwicklung von Theorien direkt aus den erhobenen Daten durch einen iterativen Prozess zum Ziel, weshalb keine theoretischen Vorannahmen zur Prüfung im Fokus stehen (Hildebrand, 2019; Przyborski & Wohlrab-Sahr, 2021). Auch sie ist mit einem sehr hohen Zeit- und Ressourcenaufwand verbunden und wenig gerichtet, was einerseits eine große Offenheit und Flexibilität ermöglicht, andererseits wenig Struktur für die Beantwortung konkreter Forschungsfragen beinhaltet. Dies liegt in ihrer Zielsetzung der Theorieentwicklung begründet, die keine oder wenig vorgefasste Hypothesen und maximale Offenheit voraussetzen (Böhm, 2019). Über die Theorien der Kognitionspsychologie lassen sich jedoch durchaus Vorannahmen in Bezug auf die hier erhobenen Daten treffen, auch wenn der Ansatz zur Operationalisierung explorativer Natur ist. Die Grounded Theory überschreitet damit die Anforderungen des Forschungsvorhabens, da eine gezielte Beantwortung konkreter Forschungsfragen unter Berücksichtigung der theoretischen Grundlage der Kognitionspsychologie im Vordergrund steht und daher verstärkt eine Strukturierung der Interviewdaten angestrebt wird. Vor diesem Hintergrund werden strukturierende Methoden, wie die qualitative Inhaltsanalyse, den rekonstruktiven Verfahren vorgezogen.

Auch unter den strukturierenden Vorgehensweisen lassen sich verschiedene Ansätze unterscheiden, wobei hier primär auf die qualitative Inhaltsanalyse eingegangen werden soll. Die qualitative Inhaltsanalyse hat dabei vorrangig das Ziel, durch kriteriengeleitete Reduktion und Kategorisierung das zu analysierende Material so aufzubereiten, dass ein repräsentatives und quantifizierbares Abbild auf Grundlage des Untersuchungsinteresses entsteht (Gläser & Laudel, 2010; Gläser-Zikuda et al., 2022; Hussy et al., 2010). Sie verbindet qualitative und quantitative Analyseschritte miteinander, da sie die qualitative Bildung und Zuordnung von Kategorien und die quantitative Auswertung der Häufigkeiten ihres Auftretens miteinander verflechtet (Döring & Bortz, 2016; Gläser & Laudel, 2010; Groeben & Rustemeyer, 1995; Hussy et al., 2010; Mayring, 2010). Die Häufigkeiten sind in der Regel ordinalskaliert, da diese in eine Rangfolge gebracht werden können, was sie in diesem Rahmen für quantitative Auswertungsmethoden öffnet (Gläser & Laudel, 2010; Kuckartz, 2016; Mayring, 2010).

Die Ansätze der qualitativen Inhaltsanalysen nach Mayring (2020) und Kuckartz (2016) sind insbesondere im deutschsprachigen Raum am weitesten verbreitet, was sich in Thematisierungen und Benennung in einschlägigen Werken, wie Flick et al. (2019), Mey und Mruck (2020) oder Reinders et al. (2022), widerspiegelt. Beide zeich-

nen sich durch eine hohe wissenschaftliche Güte und die Berücksichtigung methodologischer Gütekriterien, wie Validität, Reliabilität und Nachvollziehbarkeit, aus, weisen ein starkes theoretisches Fundament auf und sind wissenschaftlich anerkannt. Gleichzeitig haben beide Ansätze eine breite Einsetzbarkeit, können in Mixed-Method-Designs integriert werden und sind sowohl für kleine als auch große Datensätze handhabbar. Dennoch unterscheiden sich beide Ansätze in mehreren zentralen Aspekten, die für die Auswahl der Methodik entscheidend waren.

Die qualitative Inhaltsanalyse mit induktiver Kategorienbildung nach Kuckartz (2016) bietet für die Zielsetzung der Strukturierung bei gleichzeitig explorativem Charakter der Analyse ein effizientes, strukturiertes und klar definiertes Vorgehen. Aufgrund des sowohl konzeptionellen als auch empirischen Charakters dieser Arbeit liegen zwar richtungsweisende Forschungsfragen und Theoriekonstrukte vor, dennoch handelt es sich um ein exploratives Vorgehen, besonders im Hinblick auf die Operationalisierbarkeit Kritischen Denkens für die Chemiedidaktik. Mayrings (2020) Ansatz der induktiven Kategorienbildung ist stark formalisiert und linear aufgebaut, was zu einer hohen Nachvollziehbarkeit beiträgt. Trotzdem zeigt sich diese Methodik für ein exploratives Vorgehen im Falle dieser Arbeit als zu starr, da Mayring Änderungen an den Kategorien nach ihrer Definition ablehnt und nur entweder induktive oder deduktive Kategorienbildungen möglich sind. Auch Anpassungen der Formulierung von Forschungsfragen sind aufgrund des stabilen und regelgeleiteten Charakters nicht während des Prozesses möglich. Da aber vor allem bei einem explorativen Ansatz, wie bei einer Instrumententwicklung, eine gewisse Flexibilität und Dynamik notwendig ist, um auf unerwartete Muster reagieren und eine iterative Anpassung und Erweiterung der Kategorien an das Material vornehmen zu können, zeigt sich der Ansatz von Kuckartz (2016) in diesem Projekt als die geeignetste Methode. Sie strebt eine flexible Strukturierung des qualitativ erhobenen Materials an und eignet sich auch und vor allem für derart explorative Untersuchungen. Dabei betont Kuckartz (2016) ebenfalls eine mögliche Integration in einen Mixed-Method-Ansatz und definiert sowohl induktive und deduktive als auch kombinierte Vorgehensweisen, wie im Rahmen der deduktiv-induktiven Kategorienbildung. Diese angestrebte Balance zwischen methodologischer Strenge und praktischer Umsetzbarkeit bietet gerade in der Verarbeitung großer Datenmengen Vorteile und die damit einhergehende Flexibilität ist besonders im Hinblick auf die angestrebte Operationalisierung erstrebenswert. Des Weiteren integriert Kuckartz (2016) die Verwendung von Software, wie *MAXQDA*, aktiv in seinen Ansatz und betont, dass die Verwendung dieser Werkzeuge die Transparenz und Effizienz der Analyse erheblich steigern können,

was wiederum zur Güte der Forschung beiträgt. Auch Mayring hält den Einsatz solcher Software für zulässig, betont jedoch weniger ausdrücklich deren Bedeutung. In der Folge fand die Software MAXQDA aktiv Anwendung in der Auswertung der hier qualitativ erhobenen Daten, welche nach der Transkription und Aufbereitung mit ihr analysiert und kodiert[9] wurden. Das Programm unterstützt dabei qualitative Datenanalysen von Textdokumenten, aber auch von Video- und Audiodateien, mittels verschiedener Verfahren und erleichtert die Entwicklung des Kodiersystems, indem es vergebene Codes grafisch darstellt (Gläser-Zikuda et al., 2022; Hussy et al., 2010; Kuckartz, 2016; Lamnek & Krell, 2016; VERBI - Software. Consult. Sozialforschung. GmbH, 2024). Dadurch werden die gesetzten Kodierungen transparent und auch die Arbeit mit Interkodern erleichtert, da das Schreiben textstellenbezogener Memos ermöglicht wird.

Aufgrund des iterativen und flexiblen Vorgehens, der explizit vorgesehenen Möglichkeit zur Anpassung der Kategorien sowie der integrativen Nutzung von Softwaretools zur Verarbeitung großer Datenmengen wurde der Ansatz von Kuckartz (2016) als besonders geeignet erachtet. Diese Methode ermöglicht zudem die Bearbeitung explorativer Fragestellungen und zeichnet sich durch die Kombination von induktiven und deduktiven Ansätzen aus, was den spezifischen Anforderungen des vorliegenden Forschungsvorhabens entspricht. Im Folgenden wird der Ablauf der induktiven Kategorienbildung, wie sie in den Studien S1 und S2 durchgeführt wurde, detailliert dargestellt. Ferner wird erläutert, wie sich mit dem Fortschreiten des konzeptionellen Anteils der Arbeit ab Studie S3 die deduktiv-induktive Kategorienbildung nach Kuckartz als methodische Grundlage etablierte.

Die Studien S1 und S2 bildeten den initialen Kontakt mit ersten Erhebungsinstrumenten zur Operationalisierung Kritischen Denkens. Obwohl seitens der Kognitionspsychologie bereits ein theoretischer Ansatz vorhanden war, gab es in der Literatur noch kein geeignetes Instrument zur Erhebung Kritischen Denkens auf Basis einer zielgerichteten Definition für die Chemiedidaktik. Diese Defizite werden auch von Rafolt et al. (2019) in ihrem ersten und zweiten Desiderat explizit thematisiert, indem die fachbezogene Konzeptualisierung und die Erprobung der Möglichkeiten und Grenzen Kritischen Denkens im naturwissenschaftlichen Unterricht in Verbin-

[9] In der Literatur finden sich unterschiedliche Schreibweisen, wobei sowohl die deutsche Variante „Kodieren“ als auch die aus dem Englischen stammende Form „Codieren“ Verwendung findet. Zur Sicherstellung der sprachlichen Einheitlichkeit wird in dieser Arbeit die deutsche Schreibweise „Kodieren“ bevorzugt. Eine Ausnahme bildet der Begriff „Code“, der als Anglizismus beibehalten wird, da er in der qualitativen Forschung international etabliert und als Fachbegriff anerkannt ist.

dung mit SSI gefordert werden. Da die eigenen Vorannahmen zu Beginn des Projekts noch nicht weitreichend fortgeschritten waren und keine Kategorien deduktiv aus der Literatur entnommen werden konnten, wurde sich zu Beginn für die qualitative Inhaltsanalyse mit induktiver Kategorienbildung nach Kuckartz (2016) entschieden. Diese Art der Kategorienbildung erfolgt vollständig datengetrieben und die Kategorien werden aus dem Material heraus entwickelt, ohne dass theoretische Vorannahmen vorliegen müssen. Das Ziel liegt in der themenzentrierten Kategorienbildung, um inhaltliche Kerne des Materials zu identifizieren.

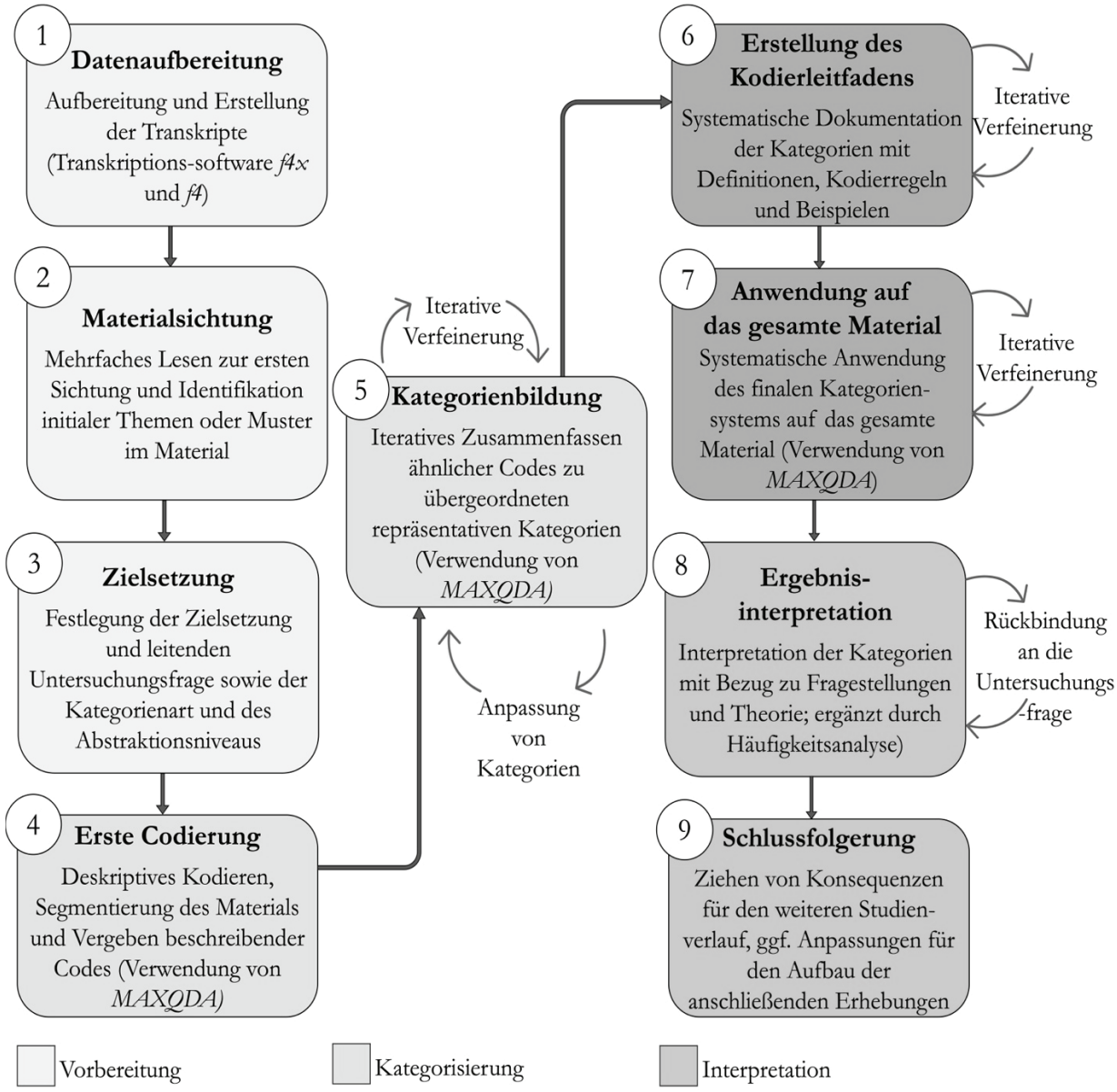


Abbildung 17: Ablauf der durchgeführten Kategorienbildung in Anlehnung an Kuckartz (2016) (eigene Darstellung)

Zur besseren Transparenz wurde in Abbildung 17 der Verlauf der Datenverarbeitung schrittweise dargestellt. Den Beginn einer jeden Auswertung der Daten mittels der

qualitativen Inhaltsanalyse nach Kuckartz (2016) bildete wie vorangehend beschrieben die Datenaufbereitung (Schritt 1), welche von einer ersten Materialsichtung (Schritt 2) begleitet wurde. Da sowohl die Interviewführung als auch die Transkripte selbst erstellt und diese mehrfach überprüft wurden, fand währenddessen bereits ein erstes Einarbeiten und die Identifikation von Themen und Mustern statt. Im dritten Schritt wurden die Zielsetzung und die leitenden Untersuchungsfragen festgelegt, welche für die Analyse handlungsleitend waren. Damit verbunden war auch die klare Definition der Auswertungseinheit, Kontexteinheit, Kodiereinheit, des Abstraktionsniveaus und der Strukturdimension sowie der Kategorienart. Trotz des explorativen Charakters der Analyse ist eine Vorabdefinition dieser Parameter ausschlaggebend für eine kriteriengeleitete und nachvollziehbare Inhaltsanalyse (Kuckartz, 2016; Mayring, 2010).

Auswertungseinheit:	Alle in einer Erhebung gewonnenen vollständigen Datensätze
Kontexteinheit:	Die Einzelelemente einer Erhebung, wie das jeweilige personenbezogene Transkript oder Teile eines Fragebogens
Kodiereinheit:	Klare, bedeutungstragende Elemente in Interviews oder den Antworten auf offene Fragen, wie Satzteile oder Wortzusammenhänge
Abstraktionsniveau:	Generalisierbare Aussagen zu Faktoren, Kriterien oder Formulierungen, die in Bezug auf die Untersuchungsfrage als relevant erachtet werden; mittleres Abstraktionsniveau
Strukturdimension:	Explizit genannte Faktoren, Kriterien oder Formulierungen sowie implizite Hinweise, welche in Bezug auf die verbundene Untersuchungsfrage als relevant erachtet wurden
Kategorienart:	Thematische Kategorien mit Bezug zu einem bestimmten Thema, Argument oder einer Denkfigur (Kuckartz, 2016)

Die Auswahl von Satzteilen und Wortzusammenhängen als Kodiereinheit eignet sich bei der qualitativen Inhaltsanalyse besonders, da diese Einheiten groß genug sind, um inhaltliche Bedeutungen zu transportieren, und gleichzeitig klein genug, um präzise Kategorisierungen zu ermöglichen. Die Fokussierung auf bedeutungstragende Elemente stellt zudem sicher, dass die Analyse sich auf relevante Inhalte konzentriert, ohne redundante oder kontextlose Satzteile zu berücksichtigen. Dennoch

wurden in wenigen Fällen auch Satzzusammenhänge kodiert, wenn diese für den kontextuellen Zusammenhang relevant waren. Das mittlere Abstraktionsniveau soll eine Balance zwischen inhaltsnaher Kodierung und konzeptueller Abstraktion ermöglichen. Ein niedrigeres Abstraktionsniveau, wie auf Einwort- oder Silbenebene, birgt die Gefahr, zu detailliert zu sein und bietet keinen effektiven Mehrwert in der hier durchgeführten Analyse. Ein höheres Abstraktionsniveau, wie auf Satz- oder Absatzebene, könnte relevante Details durch Verlust der Vielschichtigkeit des Gesagten einbüßen. Durch das mittlere Abstraktionsniveau können verschiedene Muster und Themen auch innerhalb von komplexen Satzstrukturen identifiziert werden, was gerade bei der Analyse von Transkripten mündlicher Interviews einen Vorteil bietet. Die Art der Kategorien wurde auf thematische Kategorien festgelegt, welche eine wesentliche Stärke der qualitativen Inhaltsanalyse darstellen (Kuckartz, 2016) und die Strukturierung der Daten entlang der Untersuchungsfrage auf Basis wiederkehrender Themen oder Argumente ermöglichen. Da hier vordergründig Denkperspektiven identifiziert werden sollten, die thematisch gegliedert werden können, stellte diese Kategorienart die beste Option zur Analyse der Datensätze dar. Bei der Vergabe von Codes wurde zudem bewusst die Möglichkeit offengelassen, Aussagen mehreren Kategorien zuweisen zu können. Diese Vorgehensweise liegt insbesondere in der sprachlichen und inhaltlichen Komplexität qualitativer Daten begründet (Geertz, 2019; Matt, 2019). Beispielsweise lässt sich die Aussage „Das ist das, was ich, was ich eben meinte, mit Instagram, das ist primär Meinungsmache“ (S5.I3-WS23-E-M01-IF12CG, Pos. 37) sowohl der Kategorie der Selbstreflexion von Medien als externe Einflussfaktoren (SR-A-EF-ME) als auch der Kategorie des epistemologischen Verständnisses hinsichtlich der Relevanz der Herkunft der Information im Bewertungsprozess (EP-A-B-HK) zuordnen. Trotz dieser grundsätzlichen Möglichkeit wurde angestrebt, möglichst nur einen Code pro Wortzusammenhang zu vergeben, was in seltenen Fällen zu Einwortkodierungen führte. Ein Beispiel hierfür ist die Aussage „Familie, Freunde, Medien, die ich konsumiere, eigene moralische, Moral und Wertevorstellungen, die sich mit der Zeit ergeben haben, auch Schule in einer gewissen Art und Weise.“ (S5.3-WS23-E-M01-MB14BO, Pos. 34). Diese enthält die Kategorie der internen Einflussfaktoren Moral und Werte (SR-A-IF-MW), aber auch die Reflexion externer Einflussfaktoren, wie der sozialen Norm (SR-A-EF-SN), der Medien (SR-A-EF-ME) oder der Bildung (SR-A-EF-BI). Der Code für die soziale Norm markiert dabei lediglich den Wortzusammenhang „Familie, Freunde […]“. Da der kontextuelle Inhalt trotzdem erschließbar bleibt, wurden auch solche Kleinstzusammenhänge kodiert, um Mehrfachkodierungen ganzer Phrasen möglichst zu vermeiden.

Nach der ersten Materialsichtung und dem Festlegen der grundlegenden Parameter wurden im vierten Schritt erste deskriptive Kodierungen mithilfe von MAXQDA vorgenommen (Gläser-Zikuda et al., 2022; Hussy et al., 2010; Kuckartz, 2016). Die Codes wurden dabei mit beschreibenden Codes, zum Beispiel *Fragt nach persönlichem Nutzen für sich selbst*, versehen, um die eigenen Gedankengänge im Kodierprozess festzuhalten. Erst in Schritt fünf wurden die eigentlichen Kategorien gebildet, wobei die Definition und Bezeichnung der Kategorien iterativ überarbeitet wurden und sich im Zuge der Abstraktion und gleichzeitigen Fokussierung von deskriptiven Beschreibungen hin zu Schlagwörtern entwickelten. Dabei stellte dieser Schritt keine Veränderung der Beschreibung, sondern vielmehr eine Präzisierung der Kategorien dar. So wurde beispielsweise aus *Fragt nach persönlichem Nutzen für sich selbst* im weiteren Verlauf die Kategorie *Persönliche Relevanz* (PR). Durch die iterative Anpassung wurden in diesem Schritt auch mehrere ähnliche Codes zu übergeordneten repräsentativen Kategorien zusammengefasst (Gläser-Zikuda et al., 2022). So wurden beispielsweise die Kategorien *Persönliche Relevanz* (PR) und *Moral und Werte* (MW) mit weiteren ähnlichen Kodierungen unter der Überkategorie *Introspektive Überprüfung* (SR-IU) gefasst, da diese eine innere Einsichtnahme beschreiben und auf selbstbezogene Faktoren rekurrieren. Zur Konsolidierung der Kategorien und damit einhergehend zur Förderung ihrer Trennschärfe fanden mehrere Materialdurchgänge statt (Kuckartz, 2016; Lamnek & Krell, 2016). Zusätzlich wurden in Schritt sechs die so entstandenen Kategorien systematisch durch Kodierleitfäden dokumentiert, definiert und mit ausführenden Kodierregeln und Ankerbeispielen versehen (Hammann & Jördens, 2014; Kuckartz, 2016; Lamnek & Krell, 2016; Mayring, 2010). Die Kodierleitfäden werden bei der Beschreibung der Kategorien in Kapitel 5 näher erläutert und sind im Anhang (A.3) zu finden. Dieser Schritt dient zum einen der Transparenz, zum anderen aber auch der späteren Testung der Kategorien durch die Ermittlung der Interkoderreliabilität durch eine externe Rekodierung von Teilen des Materials. Das Ziel ist die Steigerung der Intersubjektivität, welche in der qualitativen Forschung als Hauptkriterium für die Güte betrachtet werden kann (Gläser-Zikuda et al., 2022; Steinke, 2019).

Nach der Finalisierung der Kategorien und des Kodierleitfadens wurde dieser in Schritt sieben auf Grundlage der festgelegten Definitionen auf das gesamte Material angewendet. Auch dieser Schritt durchlief iterative Anpassungen durch mehrfaches Durchgehen des Materials. Die so gewonnenen Ergebnisse wurden im Anschluss in Excel aufbereitet (Schritt 8) und die Häufigkeiten als quantitative Parameter dargestellt. Dies diente hauptsächlich der Darstellung von Schwerpunkten bei den Bewertungsperspektiven und der Identifikation von Tendenzen. Da die so gewonnenen

Werte als ordinalskaliert angesehen werden können (Gläser & Laudel, 2010; Kuckartz, 2016; Mayring, 2010), waren auch Vergleiche zwischen den verschiedenen Erhebungszeitpunkten mithilfe quantitativer Auswertungsmethoden, wie dem Wilcoxon-, Mann-Whitney- oder Kruskal-Wallis-Test, möglich. Diese Testverfahren werden im Kapitel 4.3.2.3 ausführlicher dargestellt.

Anschließend wurden die Analyseergebnisse in Bezug zur jeweiligen Untersuchungsfrage gesetzt und in Schritt neun daraus hervorgehende Konsequenzen oder Verbesserungen für den weiteren Erhebungs- und Studienverlauf festgehalten. Die so gewonnenen Ergebnisse dienten als neuer Ausgangspunkt für die Planung und Durchführung weiterer Erhebungen und trugen zur Verbesserung der Instrumente im übergeordneten iterativen Studienzusammenhang bei. So führte die Analyse der ersten beiden Studien zu einer Anpassung der Fragebogenstruktur und besseren Strukturierung der Fragen im Interviewleitfaden (vgl. Kap. 4.2.4).

Diese induktive Vorgehensweise bei der Analyse der Daten wurde vor allem in Studie S1 und S2 zur Anwendung gebracht. Ab Studie S3 waren die theoretischen Konzeptionen (s. Kap. 3) so weit fortgeschritten, dass die übergeordneten Kategorien aus der Theorie, genauer aus dem Triadischen Modell Kritischen Denkens in der Naturwissenschaftsdidaktik (Abbildung 13), abgeleitet werden konnten. Die untergeordneten Kategorien wurden aus den Voruntersuchungen bei Passung eingeordnet und im weiteren Verlauf induktiv ergänzt. Der in Abbildung 17 dargestellte Ablauf blieb weitestgehend erhalten, wurde aber in Schritt 3 um die theoretische Grundlage der Definition Kritischen Denkens, insbesondere die Modellierung für die Chemiedidaktik, erweitert, was in Schritt fünf zur vorangehenden Strukturierung durch die deduktiven Überkategorien *Naturwissenschaftliches Wissen* (NW), *Selbstreflexion* (SR) und *Epistemologisches Verständnis* (EP) führte.

Zur Sicherstellung der Intersubjektivität wurden ab Studie S3 Rekodierungen des Materials durch einen weiteren Interkoder durchgeführt. Dieser war ein fortgeschrittener Doktorand der Chemiedidaktik, welcher sich aufgrund seiner fachlichen Expertise, insbesondere im Hinblick auf SSI, sowie methodischer Kompetenz besonders für diese Aufgabe eignete. Die Rekodierung erfolgte stets auf Grundlage der erstellten Kodierleitfäden in einer eigenen MAXQDA-Datei und umfasste mindestens 50 % des gesamten Datenmaterials, um eine möglichst hohe Aussagekraft zu gewährleisten. Die Ergebnisse der Rekodierungen wurden anschließend mithilfe von Excel in einer Kreuztabelle gegenübergestellt, um Übereinstimmungen und Abweichungen systematisch zu analysieren. Zur quantitativen Bewertung der Interkoderreliabilität wurde anschließend der Cohens-Kappa-Koeffizient berechnet, wodurch

die Übereinstimmung zwischen den Kodierungen statistisch erfasst wurde. Dieser Koeffizient bereinigt die Betrachtung der prozentualen Übereinstimmung um die Zufallsübereinstimmung und stellt daher ein verlässlicheres Reliabilitätsmaß dar als eine rein deskriptive Betrachtung auf Basis der prozentualen Übereinstimmung (J. Cohen, 1960; Döring & Bortz, 2016). Die Kreuztabellen zeigen auf der Diagonalen die Häufigkeiten der Übereinstimmungen der beiden Koder, während abweichende Kodierungen oberhalb oder unterhalb dieser Diagonale eingetragen wurden. Die beobachtete prozentuale Übereinstimmung (P_o) wurde wie folgt berechnet:

$$P_o = \frac{Summe\ der\ Übereinstimmungen}{Gesamtanzahl\ der\ Kodierungen}$$

(Berechnung der prozentualen Übereinstimmung (P_o), Formel in Anlehnung an Cohen, 1960, S. 39; Döring & Bortz, 2016, S. 567)

Die erwartete Übereinstimmung (P_e), welche auf der Annahme basiert, dass die Übereinstimmung rein zufällig erfolgt, ergab sich im nächsten Berechnungsschritt durch folgende Formel:

$$P_e = \sum \left(\frac{Zeilensumme}{Gesamtanzahl} \times \frac{Spaltensumme}{Gesamtanzahl} \right)$$

(Berechnung der erwarteten Übereinstimmung (P_e), Formel in Anlehnung an Cohen, 1960, S. 39; Döring & Bortz, 2016, S. 568)

Auf Basis dieser beiden Werte konnte dann der Reliabilitätskoeffizient Cohens-Kappa (κ) berechnet werden:

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

(Berechnung des Reliabilitätskoeffizienten Cohens-Kappa (κ), Formel in Anlehnung an Brennan & Prediger, 1981, S. 688; Cohen, 1960, S. 40; Döring & Bortz, 2016, S. 568)

Die so berechneten Werte liegen zwischen 0 und 1 und können rekurrierend auf Landis und Koch (1977, S. 165) folgendermaßen interpretiert werden:

< 0,00	Keine bis sehr geringe Übereinstimmung
0,00 – 0,20	Geringe Übereinstimmung
0,21 – 0,40	Mäßige Übereinstimmung
0,41 – 0,60	Substanzielle Übereinstimmung
0,61 – 0,80	Hohe Übereinstimmung
0,81 – 1,00	Nahezu perfekte Übereinstimmung

Ein erreichter Wert von $\varkappa = 0{,}70$ wird dabei für die Methode der qualitativen Inhaltsanalyse als ausreichend erachtet (Mayring, 2019).

Während sich die Kategorienbildung an Kuckartz (2016) orientierte, waren beim Auswertungsprozess, insbesondere bei der Interkodierung, auch Elemente von Mayring (2010) sinngebend. Ein Beispiel hierfür ist die Durchführung von Kodierkonferenzen, um die Kodierungen zu diskutieren und zu reflektieren, da Mayring diesem Aspekt in der Interkodierung einen besonderen Stellenwert beimisst. Dafür wurden im Anschluss an die durch den Interkoder durchgeführte Rekodierung bilaterale Kodierkonferenzen abgehalten, um differente Interpretationen und Kodierungen zu besprechen (ebd.). Die Sichtweisen, Einigungen und Ablehnungen wurden in der in MAXQDA integrierten Memofunktion zu Dokumentationszwecken festgehalten. Beispielsweise wurde der Satz „Und mein Ziel es immer ist, dass wir immer besser zusammenleben können und ich niemandem auf die Füße treten will." (S5.I2-WS23-D-M01-NB13SH, Pos. 40) vom Koder zur Kategorie *Moral und Werte* (SR-A-IF-MW) kodiert, da die Formulierung als Ziel eine Wertgrundlage widerspiegelt und hier die wertgebundene Verpflichtung aufzeigt, anderen nicht zu schaden. Interkoder 1 kodierte diese Stelle als *Soziale Norm* (SR-A EF-SN), weshalb sie in der Kodierkonferenz diskutiert wurde. Da die soziale Norm und der damit verbundene Konformitätsdruck zwar impliziert werden könnte, hier aber die eigene Zielsetzung explizit im Vordergrund steht, wurde sich auf die Kodierung zu *Moral und Werte* (SR-A-IF-MW) geeinigt, die Differenz in der Kodierung aber vermerkt und bei der Berechnung der Interkoderreliabilität berücksichtigt.

Ein detaillierter Überblick über die rekodierten Stichproben und auf diese Weise ermittelte Interkoderreliabilität der Kategorien ist in den Kapiteln 5.3.3, 5.4.3 und 5.5.3 zu finden. Eine genaue Beschreibung der Kodierleitfäden und ermittelten Kategorien wird in Kapitel 5 vorgenommen.

4.3.2 Quantitative Auswertungsmethoden

Die Daten der Fragebogenerhebungen wurden über die in LimeSurvey integrierte Export-Funktion in Excel-Arbeitsblätter übertragen und die Werte dort weiterverarbeitet. Der größte Teil der deskriptiven Statistik wurde mittels des Tabellenkalkulationsprogramms Excel erarbeitet (Microsoft Corporation, 2024) und für die Durchführung komplexerer nichtparametrischer Testverfahren auf die Datenanalysesoftware SPSS zurückgegriffen (IBM Deutschland GmbH, 2024). Zur besseren Datenverarbeitung wurden die kombiniert beschrifteten gestuften Ratingskalen in reine Zahlenwerte übersetzt und inverse Items rekodiert (Moosbrugger & Kelava, 2020a).

Die in den diskret gestuften Ratingskalen (Likert-Skalen) verwendeten Items, einschließlich der SRIS nach Silvia (2021), können als quasi-metrisch eingestuft werden, da die Kriterien der Mindeststufigkeit sowie der semantisch gleichen und numerisch präzisierten Abstände erfüllt sind. Dadurch wurden erweiterte statistische Operationen bei der Datenanalyse ermöglicht, wie die Berechnung von Mittelwerten und Standardabweichungen (Urban & Mayerl, 2011; Völkl & Korb, 2018). Im Folgenden werden die unterschiedlichen Kenngrößen der deskriptiven Statistik erläutert, die zur Beschreibung der zentralen Tendenzen und Streuungen der erhobenen Daten dienen, und es wird auf die verwendeten Testverfahren zur Hypothesentestung und Instrumentenüberprüfung eingegangen.

4.3.2.1 Maße der zentralen Tendenz

Zum initialen Überblick über die zentrale Tendenz wurde der Modus, auch als Modalwert bezeichnet, dargestellt, welcher den am häufigsten auftretenden Testwert der Verteilung repräsentiert (Hussy et al., 2010; Kelava & Moosbrugger, 2020; Kronthaler, 2021; Reinders & Post, 2022; Völkl & Korb, 2018). Die Ermittlung des Wertes fand über die Funktion =MODUS.EINF in Excel statt (Nahrstedt, 2017). Zusätzlich dazu wurden zur besseren Einschätzung der Ausprägungen der Maximalwert sowie der Minimalwert mit den entsprechenden Funktionen =MAX und =MIN in Excel bestimmt (Kronthaler, 2021; Nahrstedt, 2017; Völkl & Korb, 2018). Diese Werte können bereits ab einer Nominalskala angewendet werden und bilden daher die ersten Betrachtungspunkte der deskriptiven Statistik (Hussy et al., 2010).

Der Median setzt eine Ordinalskala voraus und steht für den Wert, der bei einer geordneten Verteilung der Werte die Stichprobe hälftig teilt. Damit ist der Median der Wert, der sowohl von 50 % der Stichprobe unterschritten und erreicht als auch überschritten und erreicht wurde (Hussy et al., 2010; Kelava & Moosbrugger, 2020;

Kronthaler, 2021; Reinders & Post, 2022; Völkl & Korb, 2018). Auch dieser Wert wurde mithilfe der in Excel zur Verfügung stehenden Funktion =MEDIAN ermittelt (Cleff, 2015; Kronthaler, 2021; Nahrstedt, 2017).

Das arithmetische Mittel (*M* oder $\bar{X}$), auch als Mittelwert bezeichnet, beschreibt den Durchschnittswert, welcher durch die Summation aller Werte und Teilung durch deren Anzahl (*N*) ermittelt wird (Hussy et al., 2010; Kronthaler, 2021; Reinders & Post, 2022; Völkl & Korb, 2018). Seine Berechnung setzt ein Intervallskalenniveau voraus, kann hier aber aufgrund der verwendeten quasi-metrischen Ratingskalen berechnet werden. Der Wert x_i dient hier als Laufindex der Messwerte der einzelnen Personen *i* bei der Variable *x*.

$$M = \frac{\sum_{i=1}^{N} x_i}{N}$$

(Berechnung des arithmetischen Mittels, Formel aus Hussy et al., 2010, S. 166)

Der Mittelwert ist eine der grundlegendsten statistischen Werte für die Beschreibung der zentralen Tendenz und ein wichtiges Maß, insbesondere in der Verhaltensforschung (J. Cohen, 1988). Er enthält das Maximum der in den Daten enthaltenen Informationen und ist damit das präferierte Maß zur Bestimmung einer Verteilung (Völkl & Korb, 2018). Auch dieser Wert wurde in Excel über die Funktion =MITTELWERT direkt ermittelt und dargestellt (Cleff, 2015; Nahrstedt, 2017).

4.3.2.2 Streumaße

Aufbauend auf der Berechnung des Mittelwertes wurden die Streuungsmaße der Varianz und Standardabweichung berechnet, um die Dispersion der einzelnen Werte darzustellen (Hussy et al., 2010; Kelava & Moosbrugger, 2020; Reinders & Post, 2022). Die empirische Varianz ist die durchschnittliche quadrierte Abweichung eines jeden Wertes vom Mittelwert und verdeutlicht so die Streuung um den Mittelwert herum (Hussy et al., 2010; Kelava & Moosbrugger, 2020; Reinders & Post, 2022; Völkl & Korb, 2018). Das Quadrieren umgeht dabei ein Ausgleichen von positiven und negativen Ausreißern vom Mittelwert (Völkl & Korb, 2018).

$$Empirische\ Varianz\ s_{emp}^2 = \frac{\sum_{i=1}^{N}(x_i - M)^2}{N}$$

(Berechnung der Varianz, Formel nach Duller, 2019, S. 121 und Hussy et al., 2010, S. 167)

Die empirische Standardabweichung *s* (in manchen Werken auch sd für standard deviation, vgl. Hussy et al., 2010) ergibt sich aus der Quadratwurzel der Varianz:

$$Empirische\ Standardanweichung\ s_{emp} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - M)^2}{N}}$$

(Berechnung der Varianz, Formel nach Hussy et al., 2010, S. 167)

Die Berechnungen von Varianz und Standardabweichung unterliegen jedoch Verzerrungen, wenn es sich um Stichproben handelt und nicht alle Werte einer Verteilung bekannt sind. Dies führt dazu, dass die unkorrigierte Varianz einer Stichprobe häufig die tatsächliche Varianz der Population unterschätzt. Um sich einer repräsentativeren Darstellung zu nähern, wird daher die theoretische oder auch korrigierte Varianz und Standardabweichung mithilfe der Bessel-Korrektur berechnet, indem nicht die Stichprobengröße *N* in den Nenner gestellt wird, sondern durch $N - 1$ geteilt wird (Cleff, 2015; Duller, 2019; Nobach, 2023). Die korrigierte Berechnung von Varianz und Standardabweichungen lässt sich wie folgt ausdrücken:

$$Korrigierte\ Varianz\ s_{korr}^2 = \frac{\sum_{i=1}^{N}(x_i - M)^2}{N - 1}$$

$$Korrigierte\ Standardanweichung\ s_{korr} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - M)^2}{N - 1}}$$

(Berechnung der korrigierten Varianz und Standardabweichung, Formel in Anlehnung an Cleff, 2015, S. 57 und Duller, 2019, S. 268)

Beide Rechenoperationen wurden direkt in Excel durchgeführt, wobei die Varianz mit der Funktion =VAR.S und die Standardabweichung mit =STABW.S berechnet wurden. Diese Funktionen beziehen sich mit ihrem Suffix *S* auf die Stichprobe und berechnen sowohl die Varianz als auch die Standardabweichung unter Anwendung der Bessel-Korrektur (Kronthaler, 2021). Auf diese Weise lässt sich ein exakteres Bild der Streuungen der gewonnenen Werte erreichen, welches näher an der tatsächlichen Populationsvarianz liegt. Dies trägt zur interpretativen Genauigkeit der Streuungsmaße bei, was besonders bei statistischen Analysen kleinerer Stichproben relevant ist (Cleff, 2015).

4.3.2.3 Vergleichende Testverfahren

Für den Vergleich zwischen Pre- und Postgruppe einer Erhebung wurde ab der Studie S3 zusätzlich zur Erhebung der zentralen Tendenzen und Streumaße auch ein Test zur Überprüfung auf signifikante Unterschiede eingesetzt. Da es sich bei Pre- und Postgruppen um verbundene Stichproben handelt, bieten sich in der Regel zwei Testverfahren mit verschiedenen Voraussetzungen an. Der t-Test bei verbundenen

Stichproben empfiehlt sich, wenn große Stichproben ($N > 30$) und mindestens intervallskalierte Daten vorliegen sowie von einer Normalverteilung auszugehen ist. Im Falle der in diesem Projekt durchgeführten Erhebungen sind die Stichprobengrößen kleiner und trotz der Voraussetzungen, die eine quasi-metrische Verarbeitung der Daten zulassen, streng genommen ordinalskaliert. Obwohl der Großteil der Daten eine annähernde Normalverteilung nahelegt, weisen zudem einzelne Items Ausreißer in der Schiefe und der Kurtosis auf (Cleff, 2015). Auch wenn der t-Test als robust gegen Voraussetzungsverletzungen gilt (Bortz & Schuster, 2010; Rasch et al., 2014a), erscheint unter den gegebenen Bedingungen der Wilcoxon-Test für abhängige Stichproben als besser geeignetes Verfahren, da er als nichtparametrischer Test eine höhere Anpassungsfähigkeit bietet und zur Wahrung der Testgüte beiträgt (ebd.).

Der *Wilcoxon-Test*, auch Wilcoxon-Vorzeichen-Rang-Test, ist ein nichtparametrischer Test zum Vergleich zweier verbundener Stichproben. Er eignet sich besonders bei nicht normalverteilten, selbst sehr kleinen Stichproben und prüft, ob signifikante Unterschiede zwischen diesen bestehen (Bortz & Schuster, 2010; Kronthaler, 2021; Rasch et al., 2014b; Wilcoxon, 1945). Dies ermöglicht sowohl eine Überprüfung der Items mit gestuften Ratingskalen innerhalb der Fragebögen als auch die statistische Untersuchung der nicht intervallskalierten Ergebnisse der Kategoriensysteme. Zur Überprüfung der H_0-Hypothese, welche davon ausgeht, dass keine signifikanten Unterschiede zwischen den beiden Stichproben vorliegen, werden die Differenzen (d_i) zwischen den Messwerten vor und nach der Intervention berechnet und deren Absolutwerte in eine Rangfolge gestellt (Wilcoxon, 1945). Obwohl die Betragswerte zur Bildung der Rangfolge benutzt werden, werden die Vorzeichen der ursprünglichen Differenzen zur Beurteilung der Veränderungsrichtung berücksichtigt und zwischen seltenen und häufig auftretenden Vorzeichen unterschieden. Die Summe der Rangplätze mit den selteneren Vorzeichen wird als T bezeichnet, wohingegen die Summe der Rangplätze der häufigeren Vorzeichen als $T´$ gekennzeichnet wird (Bortz & Schuster, 2010). Bei Wertepaaren mit einer Differenz von null bleiben diese Paare unberücksichtigt und das n wird um die Anzahl der nicht berücksichtigten Wertpaare reduziert. Beide T-Werte werden miteinander verglichen, wobei n = die Anzahl an Paardifferenzen repräsentiert:

$$T + T´ = \frac{n * (n + 1)}{2}$$

(Vergleich der beiden T-Werte im Wilcoxon-Test, Formel in Anlehnung an Bortz und Schuster, 2010, S. 133)

Ein deutlicher Unterschied zwischen *T* und *T´* spricht für eine systematische Veränderung in eine Richtung, wohingegen ähnliche Werte die Nullhypothese stärken. Um die Rangsumme der Differenzen mit dem erwarteten Mittelwert zu vergleichen, wird μ_T berechnet. Je stärker der empirische Wert von μ_T abweicht, desto mehr ist von davon auszugehen, dass die Unterschiede nicht zufällig sind (ebd.).

$$\mu_T = \frac{n * (n-1)}{4}$$

(Berechnung von μT in Anlehnung an Bortz und Schuster, 2010, S. 134)

Der empirische T-Wert wird dann mit einem kritischen *T*-Wert verglichen, der in einer Tabelle bei einem gegebenen Signifikanzniveau α und der Stichprobengröße nachgeschlagen werden kann (ebd., vgl. Tabelle 15).

Tabelle 15: Kritischen Werte für den Wilcoxon-Test, aus Bortz und Schuster, 2010, zitiert nach Clauss und Ebner, 1971, ebenfalls zu finden in Rasch et al., 2014a, S. 164

	Signifikanzniveau α für einseitige Fragestellung						
	0,025	0,01	0,005		0,025	0,01	0,005
	Signifikanzniveau α für zweiseitige Fragestellung						
n	0,05	0,02	0,01	*n*	0,05	0,02	0,01
6	0			16	30	24	20
7	2	0		17	35	28	23
8	4	2	0	18	40	33	28
9	6	3	2	19	46	38	32
10	8	5	3	20	52	43	38
11	11	7	5	21	59	49	43
12	14	10	7	22	66	56	49
13	17	13	10	23	73	62	55
14	21	16	13	24	81	69	61
15	25	20	16	25	89	77	68

Zur Vereinfachung der Prüfung der Hypothese H_0 mithilfe des Wilcoxon-Tests wurde auf die Datenanalysesoftware SPSS zurückgegriffen. Unter der Auswahl *Analysieren → Nicht parametrische Tests → Verbundene Stichproben* wurde der Wilcoxon-Test für gepaarte Stichproben ausgewählt und durch die Software die *asymptotische Signifikanz (zweiseitig)*[10], die *exakte Signifikanz (zweiseitig)*, die *exakte Signifikanz (einseitig)*, die *Punktwahrscheinlichkeit* und der *Z-Wert* ausgegeben (Janssen & Laatz, 2017). Für die Auswertung des Wilcoxon-Tests sind im Kontext der in den Studien dieser Arbeit gewonnenen Daten hauptsächlich die exakte Signifikanz (zweiseitig) und der Z-Wert von Bedeutung. Die exakte zweiseitige Signifikanz beschreibt die Wahrscheinlichkeit

[10] Bei den Schreibweisen handelt es sich um die konkreten Eigennamen der durch SPSS ausgegebenen Werte, welche daher inklusive ihrer Großschreibung beibehalten wurden.

für den Testwert ohne Annahme einer Normalverteilung und wird vor allem bei kleinen Stichproben verwendet. Sie beruht nicht auf asymptotischen Annahmen, sondern auf genaueren Berechnungen. Die Zweiseitigkeit bezeichnet die unvoreingenommene Prüfung auf Unterschiede in sowohl positive als auch negative Richtung, setzt also keine Annahme über die Änderungsrichtung der Daten voraus. Gerade bei den Vergleichen zwischen Pre- und Posterhebung ist zwar eine positive Tendenz zu erhoffen, kann aber nicht grundsätzlich angenommen werden, weshalb diese Zweiseitigkeitsprüfung zur Güte der Erprobung der Instrumente beiträgt. Das übliche Signifikanzniveau bei Hypothesentests dieser Art beträgt $\alpha = 0{,}05$ (Bortz & Schuster, 2010; Janssen & Laatz, 2017). Unterschreitet der Wert der exakten zweiseitigen Signifikanz diese Schwelle, so deutet das auf eine signifikante Veränderung hin und die Nullhypothese H_0 kann verworfen werden. In diesem Fall kann angenommen werden, dass der beobachtete Unterschied nicht zufällig zustande gekommen ist, sondern auf eine systematische Veränderung hinweist. Ist die exakte zweiseitige Signifikanz hingegen größer als das Signifikanzniveau, bleibt die Nullhypothese bestehen, was darauf hindeutet, dass keine signifikanten Veränderungen vorliegen (ebd.).

Der *Z*-Wert wiederum repräsentiert die standardisierte Teststatistik und ist für die Berechnung der Effektstärke anhand des Pearson-Korrelationskoeffizienten *r* relevant (Fritz et al., 2012). Diese ergibt sich aus der Division des *Z*-Wertes durch die Quadratwurzel der Stichprobengröße:

$$r = \frac{Z}{\sqrt{N}}$$

(Berechnung der Effektstärke anhand des Pearson-Korrelationskoeffizienten *r*, Formel in Anlehnung an Fritz et al., 2012, S. 12)

Die Effektstärke kann nach J. Cohen (1988) wie folgt interpretiert werden:

$r = 0{,}1$ geringer Effekt
$r = 0{,}3$ mittlerer Effekt
$r = 0{,}5$ großer Effekt

Je näher der Wert von *r* an -1 oder 1 liegt, desto stärker ist der Effekt ausgeprägt. Das Vorzeichen zeigt dabei die Richtung des Effekts an. Bezogen auf die in den Erhebungen der Studien gewonnenen Daten lässt sich auf diese Weise sowohl feststellen, ob die beobachteten Veränderungen statistisch signifikant sind als auch die

Stärke des beobachteten Effekts einschätzen. Trotz der zum Teil sehr kleinen Stichproben, besonders bei der Testung der Ergebnisse der Kategoriensysteme von $N < 10$, ist der Wilcoxon-Test dennoch für die Analyse geeignet, da er robust gegenüber Verteilungsannahmen ist und selbst bei derlei Stichprobengrößen verlässliche Ergebnisse liefert (Rasch et al., 2014b). Er stellt damit eine sinnvolle Wahl dar, um die Signifikanz und die Effektstärke der Veränderungen in den gewonnenen Daten zu bewerten.

Ein weiteres hier zur Anwendung gekommenes vergleichendes Testverfahren ist der *Mann-Whitney-U-Test*. Er dient der Analyse von Unterschieden zwischen zwei unabhängigen Gruppen und wird auch als Wilcoxon-Rangsummentest bezeichnet. Als nichtparametrischer Test findet er insbesondere dann Anwendung, wenn die Voraussetzungen für parametrische Tests durch Abweichungen von der Normalverteilung oder ordinalskalierte Daten nicht erfüllt sind (Bortz & Schuster, 2010; Rasch et al., 2014b). Der Mann-Whitney-U-Test eignet sich besonders für kleine Stichprobengrößen und ist robust gegenüber Ausreißern, da er auf Rangordnungen basiert und keine Verteilungsannahmen über die Grundgesamtheit trifft (H. B. Mann & Whitney, 1947). Er wurde hauptsächlich für die Analyse der Bachelor-Studierenden in Studie S3 im Vergleich zu den Master-Studierenden aus Studie S5 eingesetzt, da sich beide Studien nicht nur bei den Teilnehmenden, sondern auch in den verwendeten Kontexten unterschieden. Ziel der Untersuchung war es, mögliche Unterschiede sowohl in den Ausgangsvoraussetzungen der Pre-Tests als auch in den Ergebnissen der Post-Tests zu identifizieren. Bei den in den Studien erhobenen Daten handelte es sich um mindestens ordinalskalierte Daten und beide Gruppen konnten hinsichtlich der Teilnehmenden als unabhängig betrachtet werden, weshalb sich der Mann-Whitney-U-Test optimal eignete, um statistisch belastbare Ergebnisse zu erzielen.

Die H_0-Hypothese des U-Tests geht von keiner signifikanten Differenz zwischen beiden Gruppen aus (Rasch et al., 2014b). Zur Testung dieser Hypothese werden zunächst alle Werte beider Gruppen in eine Rangfolge gebracht und anschließend wieder separat die Rangsummen der Werte einer jeden Gruppe bestimmt. Für gleiche Messwerte, die bei der Rangbildung aufeinanderfolgen, wird ein Mittelrang gebildet. Liegen beispielsweise zwei gleiche Messwerte auf Rang 5 und 6, ergibt sich der Mittelrang 5,5 für jeden Wert. Die Rangsumme von Gruppe 1 wird als T_1 und die Rangsumme von Gruppe 2 als T_2 bezeichnet. Beide Gruppen sind über die folgende Beziehung miteinander verbunden:

$$T_1 + T_2 = \frac{n * (n+1)}{2} \; (n = n_1 + n_2)$$

(Beziehung der beiden T-Werte der Gruppen nach der Rangsummenbildung des Mann-Whitney-U-Tests, n_x = Stichprobengröße der Gruppen, Formel in Anlehnung an Bortz & Schuster, 2010, S. 131)

Die Prüfgrößen U und U´ ergeben sich aus der Auszählung der Rangplatzdifferenzen zwischen beiden Gruppen. Für jeden Wert der Gruppe 1 wird ermittelt, wie viele Werte der Gruppe 2 eine höhere Rangsumme aufweisen, wobei die Summe dieser Rangplatzüberschreitungen als *U*-Wert bezeichnet wird. Auf die vergleichbare Weise wird der U´-Wert ermittelt, welcher jedoch die Rangplatzunterschreitungen beschreibt (Bortz & Schuster, 2010; Rasch et al., 2014b). Die *U*- und *U´*-Werte können über folgende Formel berechnet werden:

$$U = n_1 * n_2 + \frac{(n_1 * n_1 + 1)}{2} - T_1$$

$$U' = n_1 * n_2 + \frac{(n_2 * n_2 + 1)}{2} - T_2$$

(Berechnung der Prüfgrößen U und U´des Mann-Whitney-U-Tests, Formeln in Anlehnung an Bortz & Schuster, 2010, S. 131; Rasch et al., 2014, S. 97)

Da sich die Werte von *U* und *U´* komplementär ergänzen, führen größere Unterschiede zwischen den beiden Gruppen auch zu größeren Unterschieden zwischen den Werten von *U* und *U´*. Die Nullhypothese, dass keine Unterschiede vorliegen, kann also folgendermaßen beschrieben werden:

$$H_0: U = U'$$

(Nullhypothese des Mann-Whitney-U-Tests, Formel in Anlehnung an Rasch et al., 2014, S. 98)

Der unter dieser Annahme erwartete mittlere *U*-Wert μ_U wird daraufhin wie folgt berechnet:

$$\mu_U = \frac{n_1 * n_2}{2}$$

(Erwarteter Mittelwert μ_U, Formel in Anlehnung an Bortz & Schuster, 2010, S. 131; Rasch et al., 2014b, S. 98)

Ein starkes Abweichen des empirischen *U*-Wertes vom erwarteten *U*-Wert spricht für eine Ablehnung der Nullhypothese. Zu Überprüfung wird bei kleineren Stichproben der empirische *U*-Wert mit dem kritischen *U*-Wert verglichen, welcher einschlägigen Werken entnommen werden kann (Tabelle 16).

Tabelle 16: Kritischen Werte bei zweiseitigem Test bei $\alpha = 0{,}05$, aus Bortz & Schuster, 2010, S. 599; ebenfalls zu finden in Rasch et al., 2014a, S. 162

	n_2											
n_1	9	10	11	12	13	14	15	16	17	18	19	20
1												
2	0	0	0	1	1	1	1	1	2	2	2	2
3	2	3	3	4	4	5	5	6	6	7	7	8
4	4	5	6	7	8	9	10	11	11	12	13	13
5	7	8	9	11	12	13	14	15	17	18	19	20
6	10	11	13	14	16	17	19	21	22	24	25	27
7	12	14	16	18	20	22	24	26	28	30	32	34
8	15	17	19	22	24	26	29	31	34	36	38	41
9	17	20	23	26	28	31	34	37	39	42	45	48
10	20	23	26	29	33	36	39	42	45	48	52	55
11	23	26	30	33	37	40	44	47	51	55	58	62
12	26	29	33	37	41	45	49	53	57	61	65	69
13	28	33	37	41	45	50	54	59	63	67	72	76
14	31	36	40	45	50	55	59	64	67	74	78	83
15	34	39	44	49	54	59	64	70	75	80	85	90
16	37	42	47	53	59	64	70	75	81	86	92	98
17	39	45	51	57	63	67	75	81	87	93	99	105
18	42	48	55	61	67	74	80	86	93	99	106	112
19	45	52	58	65	72	78	85	92	99	106	113	119
20	48	55	62	69	76	83	90	98	105	112	119	127

Auch hier wurde zur Vereinfachung der Datenanalyse auf SPSS zurückgegriffen und unter *Analysieren* → *Nicht parametrische Tests* → *Zwei unabhängige Stichproben* der Mann-Whitney-U-Test ausgewählt. Vergleichbar zum weiteren Vorgehen beim Wilcoxon-Test wurde auch hier die exakte zweiseitige Signifikanz zur Prüfung herangezogen, da keine Vorannahmen über die Veränderungsrichtung getroffen werden sollten und eine asymptotische Annäherung aufgrund der verhältnismäßig kleinen Stichprobengröße als unzureichend erachtet wurde (Janssen & Laatz, 2017). Auch für den Mann-Whitney-U-Test wurde analog zum Wilcoxon-Test die Effektstärke durch den Pearson-Korrelationskoeffizienten *r* berechnet und nach J. Cohen (1988) interpretiert.

Zum Vergleich gleicher Fragebogen-Elemente zwischen den Studien S2 bis S5 wurde aufgrund der zunehmenden Komplexität der Analyse als letztes vergleichendes Testverfahren der Kruskal-Wallis-Test verwendet. Da der Mann-Whitney-U-Test jeweils nur paarweise Vergleiche ermöglicht, hätte die Analyse der vier Studien insgesamt sechs separate Vergleiche erfordert. Dieses Vorgehen wäre nicht nur zeitaufwendig gewesen, sondern erhöht zudem das Risiko einer α-Fehler-Kumulierung, da jeder zusätzliche Test die Wahrscheinlichkeit für ein falsch-positives Ergebnis steigert (Döring & Bortz, 2016; Rasch et al., 2014b). Der Kruskal-Wallis-Test bietet hier eine effiziente Alternative, da er mehr als zwei unabhängige Gruppen gleichzeitig verglei-

chen kann. Durch das rangbasierte Vorgehen des Tests und den Verzicht auf Annahmen über die Verteilung der Daten eignet er sich besonders für die in den Studien erhobenen ordinalskalierten Daten (Kruskal & Wallis, 1952). Auf diese Weise war es möglich, alle vier Gruppen gleichzeitig auf Unterschiede in den zentralen Tendenzen zu prüfen, ohne separate paarweise Vergleiche durchführen zu müssen.

Da der Kruskal-Wallis-Test als Erweiterung des Mann-Whitney-*U*-Tests angesehen werden kann, verläuft auch die Rang- und Rangsummenbildung äquivalent zum *U*-Test (Rasch et al., 2014b). Jedoch werden hier im Anschluss die Rangsummen der Gruppen quadriert, durch die Größe der Stichprobe der jeweiligen Gruppe geteilt und aufsummiert:

$$\sum_{i=1}^{N} \frac{T_i^2}{n_i} = \frac{T_1^2}{n_1} + \frac{T_2^2}{2} + \cdots + \frac{T_k^2}{n_k}$$

(Berechnung der Summe der Quotienten aus den quadrierten Rangsummen und Stichprobengrößen mit Platzhalter *k* für die Anzahl an Gruppen[11], Formel in Anlehnung an Rasch et al., 2014b, S. 106)

Zusammen mit der Größe der Gesamtstichprobe aller betrachteten Gruppen ergibt sich der Kennwert *H*:

$$H = \left[\frac{12}{N * (N+1)}\right] * \left[\sum_{i=1}^{N} \frac{T_i^2}{n_i}\right] - 3 * (N+1)$$

(Berechnung des H-Wertes mit der Gesamtzahl der Beobachtungen *N*, Formel in Anlehnung an Kruskal & Wallis, 1952, S. 586; B. H. Cohen, 2001, S. 681; Rasch et al., 2014b, S. 107; Dodge, 2010, S. 288)

Da die Gruppengröße in keiner der Gruppen $n_i < 5$ ist, kann von einer Annäherung des *H*-Wertes an die Chi-Quadrat-Verteilung (χ^2) ausgegangen werden, wobei der Freiheitsgrad mit $df = k - 1$ beschrieben wird (B. H. Cohen, 2001). Der *H*-Wert wird dann unter Berücksichtigung der hier erfüllten Voraussetzungen mit dem χ^2-Wert verglichen, welcher in den Tabellen statistischer Werke auf Grundlage der Freiheitsgrade und des *H*-Wertes abgelesen werden kann (Tabelle 17). Liegt der *H*-Wert über dem kritischen Wert von χ^2, so kann die Nullhypothese, dass keine Unterschiede

[11] In dieser Arbeit wird das Kürzel *k* zur Bezeichnung der Anzahl der Gruppen verwendet, was in der Statistik, insbesondere im Bereich nichtparametrischer Tests, eine weit verbreitete Konvention darstellt, wie bspw. in Bortz & Schuster (2010). Während andere Werke, wie Rasch et al. (2014a), für dieselbe Größe das Kürzel *p* nutzen, wird *k* in vielen Standardwerken als Abkürzung für ‚Kategorien' oder ‚Gruppen' verwendet, insbesondere bei Verfahren wie dem Kruskal-Wallis-Test oder der Chi-Quadrat-Analyse.

vorliegen, verworfen werden und das Ergebnis ist als signifikant zu bewerten (Dodge, 2010; Rasch et al., 2014b).

Tabelle 17: Chi-Quadrat-Verteilung für den Kurskal-Wallis-Test, aus Rasch et al., 2014a, S. 165; zitiert nach Hays & Winkler, 1970, S. 604–605

Fläche *df*	0,005	0,01	0,025	0,05	0,1	0,250	0,500
1			0,001	0,004	0,016	0,101	0,455
2	0,010	0,020	0,051	0,103	0,211	0,575	1,386
3	0,072	0,115	0,216	0,352	0,584	1,212	2,366
4	0,207	0,297	0,484	0,711	1,064	1,923	3,357
5	0,412	0,554	0,831	1,145	1,610	2,675	4,351
6	0,676	0,872	1,237	1,635	2,204	3,455	5,348
7	0,989	1,239	1,690	2,167	2,833	4,250	6,346
8	1,344	1,646	2,180	2,733	3,490	5,071	7,344
9	1,735	2,088	2,700	3,325	4,168	5,899	8,343
10	2,156	2,558	3,247	3,940	4,865	6,737	9,342
11	2,603	3,053	3,816	4,575	5,578	7,584	10,341
12	3,074	3,571	4,404	5,226	6,304	8,438	11,340
13	3,565	4,107	5,009	5,892	7,042	9,299	12,340
14	4,075	4,660	5,629	6,571	7,790	10,165	13,339
15	4,601	5,229	6,262	7,261	8,547	11,037	14,339
16	5,142	5,812	6,908	7,962	9,312	11,912	15,339
17	5,697	6,408	7,564	8,672	10,085	12,792	16,338
18	6,265	7,015	8,231	9,390	10,865	13,675	17,338
19	6,844	7,633	8,907	10,117	11,651	14,562	18,338
20	7,434	8,260	9,591	10,851	12,443	15,452	19,337

Zur Vereinfachung der Berechnung wurde diese ebenfalls mithilfe von SPSS durchgeführt. Innerhalb des Programms wurde der Kruskal-Wallis-Test unter *Analysieren* → *Nicht parametrische Tests* → *K unabhängige Stichproben* ausgewählt und die zu vergleichenden Daten markiert. Neben dem Kruskal-Wallis-*H*-Wert wurden auch der Freiheitsgrad *df* und die asymptotische Signifikanz ausgegeben, da die Stichprobe in Summe bei $N > 30$ lag. Auf Grundlage dieser Werte wurde Eta^2 und der *f*-Wert berechnet, um eine Aussage über die Stärke eventueller Effekte treffen zu können. Dabei werden beide Werte wie folgt berechnet:

$$\eta_H^2 = \frac{H - k + 1}{n - k}$$

(Berechnung von Eta² (η^2) mit der Anzahl der verglichenen Gruppen *k* und der Gesamtzahl der gemachten Beobachtungen *n*, Formel nach Tomczak & Tomczak, 2014, S. 24, zitiert nach B. H. Cohen, 2008)

$$f = \sqrt{\frac{\eta^2}{1 - \eta^2}}$$

(Berechnung der Effektstärke f, basierend auf η^2, Formel in Anlehnung an J. Cohen, 1988, S. 284)

Diese Effektstärke wurde dann basierend auf J. Cohen (1988, 1992) interpretiert:

$f = 0{,}10$	geringe Effektstärke
$f = 0{,}25$	mittlere Effektstärke
$f = 0{,}40$	starke Effektstärke

Die erreichten Werte der Tests werden detailliert in Kapitel 5 berichtet und bezogen auf die einzelnen Studien umfassend vorgestellt.

5 Beschreibung der Studien

Aufgrund des stark explorativen Charakters zur Annäherung an eine Operationalisierung des Konstrukts des Kritischen Denkens und dem iterativen Aufbau der Studien ist eine ebenso iterative detaillierte Darstellung der einzelnen Studien erforderlich. Um die Komplexität und Flexibilität des Forschungsansatzes sowie die jeweiligen methodischen Entscheidungen und deren Implikationen für das Forschungsprojekt nachvollziehbar zu machen, werden daher in diesem Kapitel die Studien separat und aufeinander aufbauend behandelt. Dabei wird auf die Konzeption, Durchführung und Auswertung der jeweiligen Forschungsvorhaben eingegangen sowie auf die Erkenntnisse, die in den darauffolgenden Erhebungen berücksichtigt wurden. Zur Rekapitulation werden an dieser Stelle die in Kapitel 4.1 formulierten und der Untersuchung zugrunde liegenden Forschungsfragen erneut aufgeführt. Dabei wurde bereits in Kapitel 4.1 zwischen einer konzeptionellen und einer empirischen Forschungsfrage unterschieden, um die zwei Fokuspunkte des Projekts klar zu kennzeichnen. Die empirische Forschungsfrage, die sich explizit auf die Datenerhebung und -auswertung im Rahmen des Projekts bezieht, steht an dieser Stelle im Fokus.

Empirische Forschungsfrage:

Wie kann Kritisches Denken als Konstrukt im Kontext der Chemiedidaktik operationalisiert werden?

Die als Leitlinie dienenden Untersuchungsfragen wurden wie folgt formuliert:

1. Welche verschiedenen Perspektiven nehmen angehende Chemielehrkräfte bei der Bewertung von gesellschaftlichen Fragestellungen mit naturwissenschaftlichem Bezug (SSI) ein?
2. Welches Verständnis haben angehende Chemielehrkräfte von der Begrifflichkeit des Kritischen Denkens?
3. Inwieweit findet Selbstreflexion bei angehenden Chemielehrkräften bei der Bewertung von gesellschaftlichen Fragestellungen mit naturwissenschaftlichem Bezug statt?

Die in den einzelnen Studien behandelten Untersuchungsfragen dienten als Teilziele bei der Beantwortung der hier aufgeführten übergeordneten Untersuchungsfragen und ergänzten sich zu einer umfassenden Perspektive auf die Forschungsfrage im Kontext der Leitfrage.

5.1 Studie S1

Die Studie S1 bildete die erste Vorstudie des hier vorgestellten Projekts. Zu diesem Zeitpunkt lag noch keine präzise Konzeption der Definitionen vor, weshalb zunächst die Ausgangsvoraussetzungen der Zielgruppe im Fokus standen.

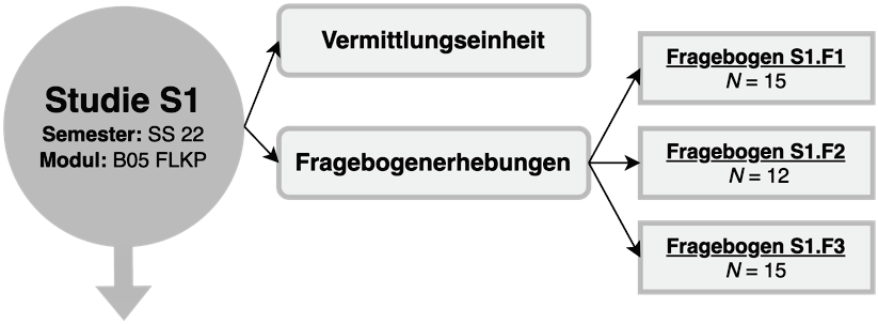


Abbildung 18: Übersicht über den Aufbau und die Erhebungsparameter der Studie S1

Die Untersuchung nahm dabei Chemielehramtsstudierende der Universität zu Köln in den Blick, denen in ihrer zukünftigen Rolle als Lehrende und Bildungsmultiplikator*innen eine besondere Verantwortung zukommt (Reiners, 2022). Da bei dieser Zielgruppe von einem grundlegenden Fachwissen ausgegangen werden konnte, rückte der Fokus der Erhebung stärker auf eingenommene Bewertungsperspektiven und die affektive Ebene. Die Ersterhebung S1 diente primär der Gewinnung erster Daten zur Orientierung und zur Erfassung der Ausgangslage der Studierenden als Grundlage für die weiteren Untersuchungen. Hierfür wurden drei Fragebogenbefragungen im *Bachelorseminar zu fachbezogenen Lern- und Kommunikationsprozessen (FLKP) (B5 HR/GG B3 SP)* durchgeführt, welches im Sommersemester 2022 vom Untersuchenden als Lehrperson gestaltet wurde (vgl. Kap. 4.2.2). Die unterschiedlichen Stichprobengrößen der einzelnen Fragebogenerhebungen trotz der Erhebung im selben Seminar resultierten daraus, dass nicht alle Studierenden durchgehend anwesend waren oder vollständige Datensätze einreichten (Abbildung 18).

Die Erhebung mittels Fragebogen bot sich vor allem aus Gründen der Praktikabilität an, da auf diese Weise eine größere Stichprobe initial befragt werden konnte, was zu einem vollständigeren Gesamtbild zur späteren Ausschärfung der Instrumente beitrug. Während der erste Fragebogen S1.F1 einen weniger gerichteten Ansatz hinsichtlich der expliziten chemischen Kontextualisierung verfolgte, um einen alltagsnahen Bewertungsprozess zu untersuchen, explizierte die Erhebung S1.F2 die fachliche Ausrichtung der dort eingesetzten Kontexte. Der abschließende Fragebogen S1.F3 erfasste letztlich konkrete Kriterien und Strategien bei der Bewertung natur-

wissenschaftlicher Informationen in Medien sowie das Verständnis von naturwissenschaftlichen Arbeits- und Denkweisen und Kritischem Denken. In dieser Fragebogenerhebung kam auch erstmals der CRT von Frederick (2005) zum Einsatz. Teil dieser ersten Studie war außerdem die Ersterprobung einer expliziten Vermittlungseinheit mit dem Schwerpunkt auf kognitiven Verzerrungen und Kritischem Denken.

Trotz des Charakters einer Vorstudie wird aufgrund der aufschlussreichen Ergebnisse, welche die Grundlage für das weitere Vorgehen legten, detaillierter auf die Durchführung und Erhebung eingegangen.

5.1.1 S1.F1 - Fragengenese 1

Im Folgenden werden die Erhebungsmethoden sowie die Auswertungsinstrumente und Ergebnisse der Fragebogenstudie S1.F1 dargestellt. Insgesamt nahmen $N = 15$ Studierende an dieser Fragebogenbefragung teil (vgl. Kap. 4.2.2). Die Befragung wurde zu Beginn des Seminars unmittelbar nach der Einführung in der ersten Sitzung durchgeführt, um einen unbeeinflussten Eindruck der Ausgangsvoraussetzungen zu gewinnen.

Unter der Annahme, dass sich das Denken im Allgemeinen und die Selbstreflexion im Besonderen durch die Fragen ausdrückt, die der Mensch in der Lage ist, sich selbst zu stellen, wurde eine Fragengenese mit den Teilnehmenden erprobt. Ziel war die Untersuchung, welche Aspekte sich für Studierende des Chemielehramts bei der Bewertung eines Kontextes mit naturwissenschaftlichem Bezug als besonders relevant darstellen und inwieweit Aspekte der Selbstreflexion Einzug finden. Die Fragengenese wurde mittels eines Fragebogens mit offenem Fragenformat über die Plattform LimeSurvey durchgeführt (s. Kap. 4.2.3). Dafür wurde zunächst ein Setting festgelegt, welches die Teilnehmenden in eine realistische Alltagssituation versetzen sollte. Reagierend auf den Kontext sollten diese Fragen formulieren, die darauf abzielen, für sich persönlich relevante Aspekte zu erfragen. Diese Fragen wurden im Anschluss von den Proband*innen nach Relevanz für die Entscheidung sortiert oder als gleichwertig angegeben. Der in S1.F1 (Studie 1, Fragebogen 1) verwendete Kontext stellte eine Entscheidung bei einem Handyneukauf nach (vgl. Kap. 4.2.6.1), wobei auch überarbeitete Geräte neben Neugeräten als gangbare Alternative formuliert wurden. Aus chemischer Sicht wären Fragen nach den Rohstoffen für die Herstellung eines Neugerätes oder der Einsparung von Rohstoffen denkbar gewesen, ebenso wie nach dem Ablauf der Überarbeitung der Neugeräte. Auf selbstreflexiver

Ebene stellten die Reflexion der Notwendigkeit eines Neugeräts oder der eigenen Werte beim Handykauf eine erwartbare Reaktion dar.

Tabelle 18: S1.F1 - Fragestellungen zur Fragengenese

Ausgangs-situation	Angenommen, Ihr aktuelles Handy wäre kaputt und nicht mehr zu retten. Sie stehen nun vor der Entscheidung, auf welche Weise Sie sich ein neues Gerät zulegen. Ihnen stehen verschiedene Kaufoptionen zur Verfügung, darunter auch gebrauchte oder überarbeitete Handys, wie von Rebuy, Backmarket u. v. m..
Frage 1	Aufgabe: Versuchen Sie, die Aspekte, die für Sie relevant sind, um eine Kaufentscheidung zu treffen, als Fragen zu formulieren, die Sie sich selbst stellen. (Formulieren Sie so viele Fragen, wie Sie für nötig erachten. Die Felder unten dienen lediglich als Platzhalter, müssen also nicht alle ausgefüllt werden.)
Frage 2	Frage: Sind alle Fragen für Sie für Ihre Entscheidungsfindung gleich bedeutsam oder gibt es für Sie Abstufungen in ausschlaggebende und weniger relevante Fragen?
Frage 3	Aufgabe: Bitte ordnen Sie Ihre Fragen nach Relevanz für Ihre Entscheidungsfindung von ausschlaggebend (oben) bis weniger relevant (unten).

Die genauen Fragestellungen des Fragebogens S1.F1 sind in Tabelle 18 dargestellt. Neben der Formulierung von Fragen im Kontext des vorgegebenen Themas sollte im weiteren Verlauf des Fragebogens, wenn möglich, auch eine Priorisierung der Fragen vorgenommen werden, um die Wertigkeit der einzelnen Aspekte transparent zu machen. Während Frage 1 als zentrale Frage des Fragebogens offen gestellt wurde und bis zu 20 formulierte Fragen in dafür vorgesehenen Kästchen eingetragen werden konnten, wurde in Frage 2 dichotom die Möglichkeit einer Priorisierung erfragt. Wurde angegeben, dass eine Abstufung der Priorität möglich ist, so wurden in Frage 3 die zuvor eingegebenen Fragen aus Frage 1 angezeigt, welche dann per Drag-and-Drop in eine Reihenfolge gebracht werden konnten. Diese Priorisierung diente zur besseren Nachvollziehbarkeit, welche Aspekte einen besonders starken oder schwachen Einfluss auf die Entscheidungen in derartigen Situationen haben, und zur Ableitung, ob stärker priorisierte Faktoren vornehmlich rational oder affektiv begründet sind.

5.1.1.1 Auswertung

Die Auswertung der Ergebnisse erfolgte unter folgender Untersuchungsfrage:

> Untersuchungsfrage: Welche Aspekte sind für Studierende des Chemielehramts bei der Bewertung eines Kontextes mit naturwissenschaftlichem Bezug relevant?

Die Ergebnisse der ausgefüllten Fragebögen wurden von LimeSurvey in Excel exportiert und für eine Kodierung in MAXQDA vorbereitet, wo sie qualitativ mittels der induktiven Inhaltsanalyse nach Kuckartz (2016) kodiert wurden (s. Kap. 4.3.1.).

Insgesamt wurden drei Überkategorien ermittelt, die in Tabelle 19 aufgeführt sind.

Tabelle 19: S1.F1 - Übergeordnete Kategorien der Fragengenese 1

Kriterialer Ansatz (KA)	Der *kriteriale Ansatz* (KA)[12] fasst alle Fragen zusammen, welche primär objektive Faktoren fokussieren. Darunter fallen beispielsweise Verständnisfragen zum Ablauf von Vorgängen oder die Frage nach Alternativen.
Reflexiver Ansatz (RA)	Der *reflexive Ansatz* (RA) fasst alle Fragen zusammen, welche primär reflexiver Natur sind und die eigenen Beweggründe oder Ansichten beleuchten oder die Konsequenzen des eigenen Handelns reflektieren. Darunter fallen beispielsweise die Reflexion der eigenen Bedürfnisse oder der Konsequenzen des Handelns.
Subjektiver Ansatz (SA)	Der *subjektive Ansatz* (SA) fasst alle Fragen zusammen, welche primär emotional oder heuristisch motiviert sind. Darunter fallen beispielsweise Fragen, die den persönlichen Vorteil zum Ziel haben oder das Bestreben für sich persönlich die beste Entscheidung zu treffen zum Ausdruck bringen.

Die Dreiteilung in *kriteriale*, *reflexive* und *subjektive Ansätze* bietet dabei eine differenzierte Analyseebene, um die Bewertung naturwissenschaftlicher Kontexte sowohl auf objektive als auch auf subjektive und reflexive Dimensionen zu untersuchen. Jede dieser drei übergeordneten Kategorien wurde zur Ausdifferenzierung in weitere untergeordnete Kategorien unterteilt.

Der *kriteriale Ansatz* (KA) beinhaltete insgesamt vier untergeordnete Kategorien mit Fokus auf externe und primär objektiv orientierte Faktoren, wie Nachfragen zum Einschätzen der Glaubwürdigkeit oder weiteren Optionen (Tabelle 20).

[12] Zur Einheitlichkeit und besseren Erkennbarkeit werden Kategoriennamen in dieser Arbeit kursiv gesetzt. Bei ihrer Verwendung als Eigenname erfolgt eine Großschreibung, wohingegen sie bei der Einbindung in den Fließtext grammatikalisch angepasst werden. Die Kursivierung bleibt dabei in beiden Fällen erhalten.

Tabelle 20: S1.F1 - Kategorien des *kriterialen Ansatzes* (KA)

Überkategorie	Kategorie	Definition
Kriterialer Ansatz (KA)	KA-MA: *Methodik/Abläufe*	Die Fragestellung zielt auf ein Verständnis von Abläufen und Methoden ab.
	KA-GL: *Glaubwürdigkeit*	Die Fragestellung zielt auf die Einschätzung der Glaubwürdigkeit ab.
	KA-WI: *Weiterführende Informationen*	Die Fragestellung zielt darauf ab, weitere Informationen einzuholen.
	KA-OV: *Optionen, Vergleichsmöglichkeiten oder Alternativen*	Die Fragestellung zielt darauf ab, einen Überblick über andere Möglichkeiten oder Alternativen zu erhalten.

Der *reflexive Ansatz* (RA) umfasst ebenfalls vier untergeordnete Kategorien, die jedoch schwerpunktmäßig reflexive Aspekte widerspiegeln, wie die Reflexion der eigenen Bedürfnisse oder der Auswirkungen der Entscheidung (Tabelle 21).

Tabelle 21: S1.F1 - Kategorien des *reflexiven Ansatzes* (RA)

Überkategorie	Kategorie	Definition
Reflexiver Ansatz (RA)	RA-EQ: *Eigene Werte und Ansichten*	Die Fragestellung hinterfragt die eigenen Werte und Ansichten.
	RA-KA: *Konsequenzen und Auswirkungen des Handelns*	Die Fragestellung zielt auf die Einschätzung der Konsequenzen des eigenen Handelns ab.
	RA-EB: *Eigene Bedürfnisse und Notwendigkeit*	Die Fragestellung hinterfragt die Notwendigkeit oder die eigenen Bedürfnisse.
	RA-EW: *Eigenes Wissen oder Bewertungsfähigkeit*	Die Fragestellung hinterfragt das eigene Wissen oder die Bewertungsfähigkeit im Bewertungskontext.

Der *subjektive Ansatz* (SA) sollte jene Aspekte unter sich gruppieren, die größtenteils ich-bezogen und in diesem Kontext emotional oder heuristisch motiviert waren (Tabelle 22). Auch vermeintlich objektiv anmutende Kategorien, wie *objektspezifische Details* (SA-OD)[13], fallen in diesen Bereich, wenn die damit verbundenen Fragestellungen primär auf den eigenen Vorteil ausgerichtet waren, wie beispielsweise die Frage nach dem Preis.

[13] Die Kürzel für die Kategorien basieren auf einem hierarchischen System, bei dem die jeweils übergeordnete Kategorie als Präfix dem spezifischen Kategorienkürzel vorangestellt wird. Dieses System erleichtert den Überblick und gewährleistet eine klare Zuordnung der einzelnen Kategorien.

Nach der Kodierung der Datensätze wurden die Ergebnisse aus MAXQDA in Excel aufbereitet und die absolute Häufigkeit der Kategorien pro Proband*in und in Summe aufgetragen. Der in den Ergebnistabellen angegebene Prozentsatz bezieht sich immer auf die Gesamtheit der formulierten Fragen. Zusätzlich wurde noch die Summe der Einzelzählungen bestimmt, was stellvertretend für die Summe der Teilnehmenden steht, die eine Frage dieser Kategorie formuliert haben. Mehrere Fragestellungen pro Proband*in mit derselben Kategorie führen in dieser Summation nicht zu einem höheren Score, was eine präzisere Betrachtung des Anteils der Teilnehmenden ermöglicht. Bei der Aufbereitung in Excel wurden höhere Werte farblich von niedrigeren differenziert, wobei mit zunehmender Höhe der Werte der Farbton intensiver oder dunkler dargestellt wurde.

Tabelle 22: S1.F1 - Kategorien des *subjektiven Ansatzes* (SA)

Überkategorie	Kategorie	Definition
Subjektiver Ansatz (SA)	SA-SG: *Subjektives Gefühl*	Die Fragestellung zielt darauf ab, Informationen zu erhalten, die basierend auf persönlichen Vorlieben als relevant angesehen werden.
	SA-SN: *Soziale Norm*	Die Fragestellung zielt darauf ab, einen Abgleich mit der sozialen Norm herzustellen.
	SA-IM: *Individueller Möglichkeitsraum*	Die Fragestellung zielt darauf ab, die persönlichen Möglichkeiten zu ermitteln.
	SA-OD: *Objektspezifische Details*	Die Fragestellung zielt darauf ab, objektspezifische Details zu erhalten.
	SA-PA: *Persönliche Absicherung*	Die Fragestellung zielt darauf ab, die persönliche Absicherung der Entscheidung zu gewährleisten.

5.1.1.2 Ergebnisse

Die 15 Teilnehmenden formulierten insgesamt 110 Fragen, was einem Durchschnitt von sieben Fragen pro Person entspricht. Keine*r der Proband*innen nutzte die Möglichkeit, alle zwanzig Fragen zu formulieren, und lediglich vier Personen reichten mehr als zehn Fragen ein.

Bei der Kodierung der im Kontext des Handyneukaufs formulierten 110 Fragen entfiel der größte Anteil auf die Überkategorie des *subjektiven Bewertungsansatzes* (SA), welche gleichzeitig die einzige Überkategorie darstellte, zu der jede teilnehmende Person mindestens eine Frage stellte. Insgesamt 65 Fragen, entsprechend 59,1 %, zeigten

einen subjektiv geprägten Entscheidungsschwerpunkt, wobei insbesondere *objektspezifische Details* (SA-OD) im Fokus standen (Tabelle 23). Dies spiegelt sich beispielsweise in Fragen wie „Wie gut ist die Kameraleistung der Smartphones bei Tageslicht?“ (S1.F1-S22-B3_5-IP30JE, Pos. 5) oder „Wie viel spare ich für ein gebrauchtes Handy?“ (S1.F1-S22-B3_5-BO38HB, Pos. 4) wider. Diese Fragen thematisierten weniger die Notwendigkeit des Kaufs, sondern fokussierten stärker affektiv geprägte Annehmlichkeiten des potenziellen Geräts. Auch der meist finanziell geprägte *Möglichkeitsrahmen* (SA-IM) wurde in 16,4 % der formulierten Fragen aufgegriffen, wie etwa bei der Frage „Welches Budget steht mir zur Verfügung?“ (S1.F1-S22-B3_5-IP30JE, Pos. 2).

Tabelle 23: S1.F1 - Ergebnisse der Fragebogenerhebung, Kategorisierung der Fragen, $N = 15$

Kategorie/Code		Summe Einzelzählung	Summe Mehrfachzählung	Summe Gesamt	Anteil Gesamt %
RA: Reflexiver Ansatz	RA-WA: Eigene Werte und Ansichten	1	3	27	**24,5**
	RA-KA: Konsequenzen und Auswirkungen	7	7		
	RA-NW: Eigene Bedürfnisse und Notwendigkeit/Sinnhaftigkeit	7	16		
	RA-EW: Eigenes Wissen oder Bewertungsfähigkeit	1	1		
KA: Kriterialer Ansatz	KA-MA: Methodik/Abläufe	2	3	18	**16,4**
	KA-GL: Glaubwürdigkeit	3	3		
	KA-WI: Weiterführende Informationen	3	3		
	KA-OV: Optionen, Vergleichsmöglichkeiten oder Alternativen	6	9		
SA: SubjektiverAnsatz	SA-SG: Subjektives Gefühls/Befindlichkeit	3	3	65	**59,1**
	SA-SN: Soziale Norm	2	4		
	SA-IM: Individueller Möglichkeitsraum	11	18		
	SA-OD: Objektspezifische Details/Fakten	12	27		
	SA-PA: Persönliche Absicherung	11	13		
Summe Fragen		15	110	7,3	Fragen/ Person (MW)
Summe reflexiver Fragen		9	27	1,8	
Summe kriterialer Fragen		10	18	1,2	
Summe subjektiver Fragen		15	65	4,3	

Dem *reflexiven Bewertungsansatz* (RA) konnten 24,5 % der Fragen zugeordnet werden, wobei hier primär die *Reflexion eigener Bedürfnisse* (RA-EB) im Vordergrund stand (Tabelle 23). Stellvertretend können hier Fragen wie „Wie notwendig ist der Kauf?“ (S1.F1-S22-B3_5-JO16OF, Pos. 2) oder „Will ich das neuste vom neusten oder reicht mir eine ältere Version?“ (S1.F1-S22-B3_5-FN33CB, Pos. 9) angeführt werden. Mit 6,4 % der Fragen spielten mögliche *Konsequenzen* (RA-KA) eine deutlich geringere Rolle für die Teilnehmenden. Nur sieben Proband*innen stellten je eine Frage, wie beispielsweise „Wie wirkt sich der Kauf indirekt auf die Umwelt aus?“ (S1.F1-S22-B3_5-JO15OF, Pos. 10).

Der *kriteriale Bewertungsansatz* (KA) war mit 16,4 % der Fragen am schwächsten vertreten. Hier waren Fragen stellvertretend, welche einen *Vergleich von Optionen* (KA-OV) anstrebten, wie „Welches Smartphone von welchem Anbieter ist am langlebigsten?“ (S1.F1-S22-B3_5-NB16OF, Pos. 5) oder „Welche Modelle mit welchen Funktionen stehen zur Verfügung?“ (S1.F1-S22-B3_5-CS21FU, Pos. 5). Weiterführende Informationen (KA-WI) wurden nur von drei Personen erfragt und waren nicht naturwissenschaftlich geprägt, wie das Beispiel „Wer ist der Händler?“ (S1.F1-S22-B3_5-CS21FU, Pos. 9) verdeutlicht.

Insgesamt zeigt sich eine recht eindeutige Verteilung zugunsten subjektiver Bewertungsansätze, welche fast zwei Drittel der Fragen ausmachen. Auch die durch die Teilnehmenden vorgenommene Priorisierung verstärkt diese Beobachtung, da zwei Drittel der Proband*innen subjektiv konnotierten Fragen die Priorität 1 zuordneten und über 80 % auch auf Priorität 2 affektiv geprägte Faktoren thematisieren (Tabelle 24). Dieser Trend zeichnete sich auch auf den späteren Rängen ab und verlagert sich erst ab Rang 8 zugunsten des reflexiven Ansatzes. Über 40 % der Fragen auf Priorität 1 und 2 thematisierten dabei den individuellen Möglichkeitsraum durch Fragen wie „Wie viel Geld habe ich auf dem Konto?“ (S1.F1-S22-B3_5-SP20KF, Pos. 2) oder „Habe ich ausreichend Geld für ein neues Smartphone?“ (S1.F1-S22-B3_5-NB16OF, Pos. 2). Kriteriale Fragen tauchten erst mit einer Priorität von 3 bis 5 verstärkt auf.

Fragen, welche dem *reflexiven Bewertungsansatz* (RA) zugeordnet wurden, waren auf den hinteren Rängen verstärkt vertreten, wobei Fragen nach den *Auswirkungen der Kaufentscheidung* (RA-KA), wie „Wie wirkt sich der Kauf indirekt auf die Umwelt aus?“ (S1.F1-S22-B3_5-JO16OF, Pos. 10), erst ab Rang 9 priorisiert wurden.

Tabelle 24: S1.F1 - Ergebnisse der Fragebogenerhebung, Priorisierung der Fragen, $N = 15$

Kategorie/Priorität		1	2	3	4	5	6	7	8	9	10	11	Summe
RA: Reflexiver Ansatz	RA-WA: Eigene Werte und Ansichten	1	1							1			**3**
	RA-KA: Konsequenzen und Auswirkungen		1		1			1		3	1		**7**
	RA-NW: Eigene Bedürfnisse und Notwendigkeit/Sinnhaftigkeit	3		2	1	1	1	1	3		2	2	**16**
	RA-EW: Eigenes Wissen oder Bewertungsfähgkeit				1								**1**
KA: Kriterialer Ansatz	KA-MA: Methodik/Abläufe			1	1	1							**3**
	KA-GL: Glaubwürdigkeit	1		1			1						**3**
	KA-WI: Weiterführende Informationen					1				1	1		**3**
	KA-OV: Optionen, Vergleichsmöglichkeiten oder Alternativen			2	3	2	1	1					**9**
SA: SubjektiverAnsatz	SA-SG: Subjektives Gefühl/Befindlichkeit		1			1					1		**3**
	SA-SN: Soziale Norm		1			1	1	1					**4**
	SA-IM: Individueller Möglichkeitsraum	6	7	1	1	1		1	1				**18**
	SA-OD: Objektspezifische Details/Fakten	3	3	5	3	3	3	2	2	1	1	1	**27**
	SA-PA: Persönliche Absicherung	1	1	2	3		3	1		2			**13**
Summe Fragen		15	15	14	14	11	10	8	6	8	6	3	110
Summe reflexiver Fragen		4	2	2	3	1	1	2	3	4	3	2	27
Summe kriterialer Fragen		1	0	4	4	4	2	1	0	1	1	0	18
Summe subjektiver Fragen		10	13	8	7	6	7	5	3	3	2	1	65

Dies legt im Hinblick auf die zugrundeliegende Untersuchungsfrage

> Welche Aspekte sind für Studierende des Chemielehramts bei der Bewertung eines Kontextes mit naturwissenschaftlichem Bezug relevant?

den Schluss nahe, dass primär subjektive Faktoren im Vergleich zu reflexiven oder kriterialen Ansätzen in Entscheidungsprozessen eine Rolle spielen, auch wenn diese einen naturwissenschaftlichen Bezug aufweisen. Dies kann viele Ursachen haben, wobei eine das Nichterkennen oder die Nichtbeachtung der naturwissenschaftlichen Komponente innerhalb des Kontextes sein könnte. Die anfängliche Befürchtung, dass der übergeordnete Erhebungskontext im Rahmen des Studiums die Antworten in Richtung einer verstärkt naturwissenschaftlichen Betrachtung verlagern würde,

konnte nicht bestätigt werden. Gerade aber im Hinblick auf die naturwissenschaftliche Grundbildung wäre das Nichterkennen einer naturwissenschaftlich geprägten Fragestellung problematisch, insbesondere da es sich bei den Befragten um angehende Lehrkräfte im Fach Chemie handelt. Dennoch muss bei der Interpretation der Ergebnisse hervorgehoben werden, dass es sich hier um den Vorläufer eines Werkzeuges handelt und die verwendeten Kategorien lediglich eine Vorstufe des später angestrebten, trennschärferen Kategoriensystems repräsentieren.

5.1.2 S1.F2 - Fragengenese 2

Die zweite Phase der Fragengenese wurde mit den Studierenden des Seminars am Ende der dritten Sitzung durchgeführt. Da die Themen der ersten vier Sitzungen innerhalb der Vermittlungseinheit eine andere chemiedidaktische Ausrichtung hatten (vgl. Kap. 4.2.5), war zu diesem Zeitpunkt noch keine Fokussierung auf den Schwerpunkt des Kritischen Denkens erfolgt, sodass ein Priming in diese Richtung ausgeschlossen werden konnte. Insgesamt nahmen $N = 12$ Studierende an der Erhebung teil (vgl. Kap. 4.2.2).

Auf Grundlage der Fragengenese 1, welche durch die alltagsnahe Handykaufentscheidung einen weitestgehend ungerichteten Kontext beinhaltete, wurde die Fragestellung der zweiten Fragengenese explizit naturwissenschaftlich-chemisch ausgerichtet. Dies wurde mit der Zielsetzung durchgeführt, zu untersuchen, inwieweit sich die gestellten Fragen durch die Betonung eines chemischen Bezugsrahmens im Vergleich zum ungerichteten Ansatz verändern und ob die Teilnehmenden in der Lage sind, bei der Auseinandersetzung mit einem explizit fachlichen Kontext eine stärker fachlich orientierte Perspektive einzunehmen.

Den Studierenden wurden hierfür drei Themen zur Auswahl gestellt, aus denen sie interessengeleitet wählen konnten. Die zur Verfügung gestellten Kontexte umfassten die Themen Fracking, Seltene Erden sowie die Wasserstoff-Brennstoffzelle und waren im späteren Verlauf ab Sitzung 8 auch für die abschließende Hausarbeit im Seminar wählbar. Zum Zeitpunkt der Erhebung wurden bis auf die gegebene Fragestellung (Tabelle 25) jedoch keine weiteren Angaben zu den Themen gemacht, um eine Vorbeeinflussung zu reduzieren.

Der Aufbau des Fragebogens sowie die Fragestellungen verhielten sich analog zum Aufbau der ersten Fragengenese. Allein die Auswahl der stärker fachlich ausgerichteten Settings differenzierte beide Erhebungen.

Tabelle 25: S1.F1 - Fragestellungen zur Fragengenese

Ausgangs-situation	Sie haben sich in der letzten Sitzung für eines der drei folgenden Themen entschieden: 1. Fracking – Druck für die Gesellschaft. Eine saubere Sache? 2. Seltene Erden – Kritisch in die Zukunft. Um jeden Preis mit Solarplatten und E-Autos in die Zukunft? 3. Brennstoffzellenauto – Brennt´s bei dir? Zukunft oder nur Knallgas?
Frage 1	Aufgabe: Welche Informationen bräuchten Sie, um sich aus chemischer Sicht an dem Thema beteiligen zu können? Formulieren Sie bitte Ihre Wissensziele als Fragen. (Formulieren Sie so viele Fragen, wie Sie es für nötig erachten. Die Felder unten dienen lediglich als Platzhalter, müssen also nicht alle ausgefüllt werden.)
Frage 2	Frage: Sind alle Fragen für Sie für Ihre Bewertung auf chemischer Ebene gleich bedeutsam oder gibt es für Sie Abstufungen in ausschlaggebende und weniger relevante Fragen?
Frage 3	Aufgabe: Bitte ordnen Sie Ihre Fragen nach Relevanz für Ihre Bewertung aus chemischer Perspektive von ausschlaggebend (oben) bis weniger relevant (unten).

5.1.2.1 Auswertung

Die Auswertung erfolgte ebenfalls analog zu S1.F1 und auch die Untersuchungsfrage sowie die übergeordneten Kategorien blieben bestehen. Allerdings veränderten sich in Reaktion auf die veränderte Ausrichtung der Kontexte der Umfang der untergeordneten Kategorien, welche daher für diese zweite Erhebung erweitert oder reduziert wurden.

Tabelle 26: S1.F2 - Kategorien des *kriterialen Ansatzes* (KA)

Über-kategorie	Kategorie	Definition
Kriterialer Ansatz (KA)	KA-MA: *Methodik/Abläufe*	Die Fragestellung zielt auf ein Verständnis von Abläufen und Methoden ab.
	KA-GL: *Glaubwürdigkeit*	Die Fragestellung zielt auf die Einschätzung der Glaubwürdigkeit von Informationen oder Quellen ab.
	KA-WI: *Weiterführende (fachliche Informationen)*	Die Fragestellung zielt darauf ab, weitere Informationen einzuholen.
	KA-OV: *Optionen, Vergleichsmöglichkeiten oder Alternativen*	Die Fragestellung zielt darauf ab, einen Überblick über andere Möglichkeiten oder Alternativen zu erhalten.
	KA-NU: *Nutzen*	Die Fragestellung zielt darauf ab, ein Verständnis des Nutzens zu erhalten.
	KA-HK: *Herkunft der Materialien*	Die Fragestellung zielt darauf ab, Informationen über die Herkunft von Rohstoffen und Materialien zu gewinnen.
	KA-VS: *Versorgungssicherheit*	Die Fragestellung zielt auf die Einschätzung der Versorgungssicherheit ab.

Zusätzlich zu den bereits bestehenden Kategorien des *kriterialen Bewertungsansatzes* (KA) wurden die Kategorien *Nutzen* (KA-NU), *Herkunft der Materialien* (KA-HK) sowie *Versorgungssicherheit* (VS) ergänzt (Tabelle 26).

Von den Subkategorien des *reflexiven Bewertungsansatzes* (RA) fanden sich in der stärker fachlich ausgerichteten Fragengenese nur zwei Aspekte wieder (Tabelle 27). Fragestellungen subjektiver Natur wurden nicht gestellt.

Tabelle 27: S1.F2 - Kategorien des *reflexiven Ansatzes* (RA)

Überkategorie	Kategorie	Definition
Reflexiver Ansatz (RA)	RA-KA: *Konsequenzen und Auswirkungen des Handelns*	Die Fragestellung zielt darauf ab, die möglichen Konsequenzen des eigenen Handelns einzuschätzen.
	RA-NW: *Notwendigkeit/Sinnhaftigkeit*	Die Fragestellung hinterfragt die Notwendigkeit des Handelns.

Die Darstellung der Ergebnisse erfolgte ebenfalls analog zur Erhebung S1.F1.

5.1.2.2 Ergebnisse

Als Reaktion auf die stark fachliche Ausrichtung der gewählten Settings in S1.F2 rückten, wie erwartet, ich-bezogene Fragestellungen derart in den Hintergrund, dass sie vollständig ausblieben. Auch reflexive Bewertungsansätze traten deutlich zugunsten kriterialer Perspektiven in den Hintergrund. Insgesamt formulierten die zwölf Teilnehmenden 90 Fragen, wovon 82,2 % dem *kriterialen Bewertungsansatz* (KA) zugeordnet werden konnten (Tabelle 28).

Tabelle 28: S1.F2 - Ergebnisse der Fragebogenerhebung, Kategorisierung der Fragen, *N* = 12

Kategorie/Code		Summe Einzel-zählung	Summe Mehrfach-zählung	Summe Gesamt	Anteil Gesamt %
KA: Kriterialer Ansatz	KA-NU: Nutzen	1	1	74	**82,2**
	KA-OV: Optionen, Vergleichsmöglichkeiten oder Alternativen	5	7		
	KA-EF: Effektivität	1	1		
	KA-HK: Herkunft (der Rohstoffe)	3	4		
	KA-VS: Versorgungssicherheit	1	1		
	KA-GL: Glaubwürdigkeit von Quellen oder dem Forschungsstand	2	3		
	KA-MA: Methodik/Abläufe/ Funktionsweisen	10	26		
	KA-WI: Weiterführende (fachliche) Informationen/Wissensfragen	10	31		
RA: Reflexiver Ansatz	RA-NW: Notwendigkeit/Sinnhaftigkeit	2	2	16	**17,8**
	RA-KA: Konsequenzen und Auswirkungen	8	14		
Summe Fragen		12	90	7,5	Fragen/ Person (MW)
Summe kriterialer Fragen		12	74	6,2	
Summe reflexiver Fragen		9	16	1,3	

Den Schwerpunkt bildeten dabei die Kategorien *Weiterführende Informationen* (KA-WI), mit einem Anteil von 36,7 % an den Gesamtfragen, sowie *Methodik/Abläufe* (KA-MA), die 28,9 % der Fragen ausmachten. Beispiele hierfür sind die Fragen „Welche chemischen Stoffe sind am Fracking beteiligt?“ (S1.F2-S22-B3_5-FN32CF, Pos. 3) für die Kategorie *Weiterführende Informationen* (KA-WI) und „Wie funktionieren Brennstoffzellen?“ (S1.F2-S22-B3_5-JO16OF, Pos. 3) für die Kategorie *Methodik/Abläufe* (KA-ME).

Fragen des *reflexiven Bewertungsansatzes* (RA) reduzierten sich nahezu vollständig auf die Kategorie *Konsequenzen und Auswertungen* (RA-KA), auf welche 15,6 % der Fragen entfielen. Exemplarisch sei hier die Fragestellung „Gibt es Studien zu den Wirkungen

jener Chemikalien auf Mensch/Umwelt?“ (S1.F2-S22-B3_5-NB16OF, Pos. 4) angeführt.

Auch die Priorisierung verlagerte sich stark zugunsten des kriterialen Ansatzes, was sich darin äußerte, dass alle Teilnehmenden Fragen *kriterialen Ursprungs* (KA), insbesondere Fragen zu *weiterführenden Informationen* (KA-WI) oder zu den *Abläufen der Technologien* (KA-MA), die höchste Priorität zuwiesen (Tabelle 29). Diese klare Tendenz legt nahe, dass die Teilnehmenden durchaus in der Lage waren, fachliche Fragen innerhalb von Bewertungskontexten zu formulieren. Jedoch schienen sie naturwissenschaftliche Aspekte entweder in bestimmten Bereichen zu übersehen oder ihnen keine wesentliche Bedeutung beizumessen. Ein Defizit in der grundlegenden Fähigkeit zur Fragengenese war jedoch nach der Erhebung S1.F2 nicht feststellbar.

Tabelle 29: S1.F2 - Ergebnisse der Fragebogenerhebung, Priorisierung der Fragen, $N = 12$

Kategorie/Priorität		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Summe
KA: Kriterialer Ansatz	KA-NU: Nutzen						1										**1**
	KA-OV: Optionen, Vergleichsmöglichkeiten oder Alternativen				1	2	1	1		1	1						**7**
	KA-EF: Effektivität			1													**1**
	KA-HK: Herkunft (der Rohstoffe)					1				1				1	1		**4**
	KA-VS: Versorgungssicherheit												1				**1**
	KA-GL: Glaubwürdigkeit von Quellen oder dem Forschungsstand	1	1													1	**3**
	KA-MA: Methodik/Abläufe/ Funktionsweisen	6	3	6	3		1	1	2	2	2						**26**
	KA-WI: Weiterführende (fachliche) Informationen/Wissensfragen	5	5	4	3	5	3	3	3			1	1				**33**
RA: Reflexiver	RA-NW: Notwendigkeit/Sinnhaftigkeit											1					**1**
	RA-KA: Konsequenzen und Auswirkungen		3		5	2	2	1						1			**14**
Summe Fragen		12	9	11	7	8	6	5	5	4	3	1	2	1	1	1	76
Summe kriterialer Fragen		12	9	11	7	8	6	5	5	4	3	1	2	1	1	1	76
Summe reflexiver Fragen		0	3	0	5	2	2	1	0	0	0	1	0	1	0	0	15

5.1.3 S1.F3 – Kriterien, Verständnis und CRT

Die dritte und abschließende Fragebogenerhebung der ersten Studie wurde am Ende der vierten Sitzung durchgeführt, da diese Sitzung die letzte Einheit darstellte, die keinen thematischen Bezug zum Forschungsprojekt aufwies. Nachdem die Untersuchung relevanter Bewertungsaspekte implizit durch die beiden Phasen der Fragengenese in S1.F1 und F2 erfolgt war, strebte die Erhebung S1.F3 die explizite Thematisierung und Untersuchung von Kriterien und Strategien bei der Bewertung naturwissenschaftlicher Informationen aus Medien durch angehende Chemielehrkräfte an und erhob zudem das Verständnis der Begrifflichkeiten des Kritischen Denkens und der naturwissenschaftlichen Arbeits- und Denkweisen. Als Ergänzung wurde außerdem erstmalig der CRT nach Frederick (2005) eingesetzt. Insgesamt nahmen $N = 15$ Studierende aus dem Seminar an der Fragebogenerhebung teil (vgl. Kap. 4.2.2). Die genauen Fragestellungen sind in (Tabelle 30) aufgeführt.

Tabelle 30: S1.F3 - Fragestellungen in der Fragebogenerhebung

Frage 1	Bitte tragen Sie Ihre Antwort auf folgende Frage im unten stehenden Textfeld ein: Nach welchen Kriterien bewerten Sie naturwissenschaftliche Informationen aus Medien, wie Zeitung, Fernsehen, sozialen Medien o. Ä.? (Seien Sie gerne ausführlich. Auch Beispiele können Sie gerne anbringen!)
Frage 2	Bitte tragen Sie Ihre Antwort auf folgende Frage im unten stehenden Textfeld ein: Welche Strategien wenden Sie an, wenn Sie sich eine Meinung zu einem naturwissenschaftlichen Thema bilden wollen, welches gerade gesellschaftlich relevant/aktuell ist? (Seien Sie gerne ausführlich. Auch Beispiele können Sie gerne anbringen!)
Frage 3	Bitte tragen Sie Ihre Antwort auf folgende Frage im unten stehenden Textfeld ein: Welche naturwissenschaftlichen Arbeits- und Denkweisen sind Ihnen bekannt bzw. was verstehen Sie unter naturwissenschaftlichem Arbeiten und Denken? (Seien Sie gerne ausführlich. Auch Beispiele können Sie gerne anbringen!)
Frage 4	Bitte tragen Sie Ihre Antwort auf folgende Frage im untenstehenden Textfeld ein: Was verstehen Sie unter Kritischem Denken? Was macht gerade das Kritische daran aus? Seien Sie gerne ausführlich. Auch Beispiele können Sie gerne anbringen!)
CRT 1	Ein Schläger und ein Ball kosten zusammen 1,10 €. Der Schläger kostet 1,00 € mehr als der Ball. Wie viel kostet der Ball? [Euro:]
CRT 2	Wenn 5 Maschinen 5 Minuten benötigen, um 5 Geräte herzustellen, wie lange dauert es mit 100 Maschinen 100 Geräte herzustellen? [Minuten:]
CRT 3	In einem See ist ein Feld mit Seerosen. Jeden Tag verdoppelt sich die Größe des Feldes. Wenn die Seerosen 48 Tage brauchen, um den gesamten See zu bedecken, wie lange dauert es, bis sie die Hälfte des Sees bedecken? [Tage:]

Die Fragen nach Kriterien und Strategien beim Umgang mit naturwissenschaftlichen Informationen sollten Aufschluss darüber geben, welche Aspekte die Studierenden

bei expliziter Nachfrage als besonders relevant für die Bewertung erachten und inwieweit sie dabei strukturiert und bewusst vorgehen. Obwohl Kriterien und Strategien auf den ersten Blick Ähnlichkeiten aufweisen, ist eine präzise Unterscheidung notwendig. Während Kriterien bestimmte Merkmale verkörpern, die für eine Einschätzung von Informationen als relevant erachtet werden, fassen Strategien konkrete Vorgehensweisen zusammen.

Zusätzlich dazu sollten die Verständnisfragen die Vorstellungen der Studierenden bezüglich relevanter Begrifflichkeiten offenlegen. Im Hinblick auf das zentrale Erkenntnisinteresse dieses Projekts stellte die Untersuchung der Präkonzepte zum Begriff des Kritischen Denkens einen konsequenten ersten Schritt dar. Darüber hinaus wurde auch das Verständnis naturwissenschaftlicher Arbeits- und Denkweisen ergründet, da in Anlehnung an eine vorangegangene Abschlussarbeit (Albrecht, 2021), die die Bedeutung naturwissenschaftlicher Arbeits- und Denkweisen für die Förderung Kritischen Denkens untersuchte, angenommen werden konnte, dass das Verständnis der Charakteristika der Naturwissenschaften einen relevanten Faktor für die Entwicklung Kritischen Denkens darstellt.

Zusätzlich zu den offenen Fragen wurde der Cognitive Reflection Test (CRT) von Frederick (2005) implementiert. Dieser besteht aus nur drei Fragen (s. Kap. 4.2.3, Tabelle 8), stellt dabei aber gleichzeitig einen der besten Prädiktoren für erfolgreiche Entscheidungsprozesse dar. Aufgrund der Einfachheit der Implementation und der gleichzeitigen guten Aussagekraft wurde der CRT im Fragebogen abwechselnd zu den offenen Fragen eingesetzt, wobei die Abfolge der CRT-Fragen durch LimeSurvey randomisiert wurde.

5.1.3.1 Auswertung

Die Auswertung der offenen Fragen und des CRT erfolgte unter den jeweils zugeordneten Forschungsfragen (Tabelle 31).

Tabelle 31: S1.F3 - Untersuchungsfragen zu den Fragestellungen

Frage 1	Nach welchen Kriterien bewerten angehende Chemielehrkräfte naturwissenschaftliche Informationen in Medien?
Frage 2	Welche Strategien wenden angehende Chemielehrkräfte bei der Bewertung von naturwissenschaftlichen Informationen in Medien an?
Frage 3	Welche naturwissenschaftlichen Arbeits- und Denkweisen sind angehenden Chemielehrkräften bekannt?
Frage 4	Welches Verständnis haben angehende Chemielehrkräfte von der Begrifflichkeit des Kritischen Denkens?
CRT	In welchem Maße sind angehende Chemielehrkräfte in der Lage, ihren intuitiven Erstimpuls zu inhibieren und zugunsten einer bewussten kognitiven Reflexion zu kontrollieren?

Dabei wurden die Antworten der offenen Fragen ebenfalls mittels der qualitativen Inhaltsanalyse nach Kuckartz mit induktiver Kategorienbildung ausgewertet und die absoluten Häufigkeiten der Kategorien dargestellt. Ziel war es nicht, Antworten abzufragen und zu bewerten, sondern das individuelle Verständnis der Teilnehmenden zu erfassen, was dem induktiven Charakter des Vorgehens entspricht.

Die Überkategorie der *Kriterien bei der Bewertung naturwissenschaftlicher Informationen in Medien* (KB) umfasste nach der Kodierung zehn differente Kategorien, welche in Tabelle 32 aufgeführt und definiert sind.

Tabelle 32: S1F3 - Kategorien der Kriterien bei der Bewertung naturwissenschaftlicher Informationen in Medien (KB)

Überkategorie	Kategorie	Definition
Kriterien bei der Bewertung naturwissenschaftlicher Informationen in Medien (KB)	KB-PL: *Plattform/Art der Quelle*	Die Plattform, auf der eine Quelle veröffentlicht wird, wird als relevantes Kriterium angegeben.
	KB-TR: *Transparenz/Angabe von Quellen und Daten*	Die Transparenz des Vorgehens und der Quellen wird als relevantes Kriterium angegeben.
	KB-VL: *Verlässlichkeit der Information/Quelle*	Die Verlässlichkeit der Information oder Quelle wird als relevantes Kriterium angegeben
	KB-PR: *Professionelle Aufbereitung (Schrift, Form, Präsentation)*	Die professionelle Aufbereitung der Information in Form und Präsentation wird als relevantes Kriterium angegeben.
	KB-EX: *Expertise des Autors*	Die Expertise der*s Autor*in wird als relevantes Kriterium angegeben.
	KB-EF: *Eigenes Fachwissen*	Das Übereinstimmen mit dem eigenen Fachwissen wird als relevantes Kriterium angegeben.
	KB-AK: *Aktualität*	Die Aktualität einer Quelle wird als relevantes Kriterium angegeben.
	KB-BK: *Bekanntheit der Quelle*	Die Bekanntheit einer Quelle wird als relevantes Kriterium angegeben.
	KB-VG: *Betrachtung/Vergleich mehrerer Quellen/Standpunkte*	Die Berücksichtigung mehrerer Standpunkte in Quellen wird als relevantes Kriterium angegeben.
	KB-KG: *Kongruenz weitere Quellen*	Die Kongruenz mit anderen Quellen wird als relevantes Kriterium angegeben.
	KB-RP: *Repräsentativität*	Die Repräsentativität der Information wird als relevantes Kriterium angegeben.

Die Kategorisierung der Antworten auf die Frage nach den *Strategien bei der Bewertung naturwissenschaftlicher Informationen in Medien* (SB) endete mit fünf Kategorien, wobei die *Quellen* (SB-QU) abermals durch Subkategorien ausdifferenziert wurden (Tabelle 33).

Tabelle 33: S1.F3 - Kategorien und Subkategorien von *Strategien bei der Bewertung naturwissenschaftlicher Informationen in Medien* (SB)

Überkategorie	Kategorie	Subkategorie	Definition
Strategien bei der Bewertung naturwissenschaftlicher Informationen in Medien (SB)	SB-VQ: *Nutzung verschiedener Quellen/ Standpunkte*		Die Nutzung verschiedener Quellen und Standpunkte wird als Strategie benannt.
	SB-OF: *Offenheit gegenüber neuen Informationen/ Standpunkten*		Die Offenheit gegenüber neuen Informationen oder Standpunkten wird als Strategie benannt.
	SB-ÜW: *Überprüfung des eigenen Wissensstandes*		Das Zurückgreifen auf das eigene Wissen oder die Überprüfung dieses Wissens wird als Strategie benannt.
	SB-AT: *Austausch*	SB-AT-LA: *Austausch mit anderen (Laien)*	Der Austausch mit anderen aus dem privaten Umfeld wird als Strategie benannt.
		SB-AT-EX: *Austausch mit anderen (Expert*innen)*	Der Austausch mit Personen mit Expertise wird als Strategie benannt.
	SB-QU: *Quellen*	SB-QU-UK: *Unkonkret*	Das Hinzuziehen einer unkonkreten Quelle wird als Strategie benannt.
		SB-QU-FL: *Forschungsliteratur*	Das Hinzuziehen von Forschungsliteratur/ Papern wird als Strategie benannt.
		SB-QU-LB: *Lehrwerke/ Bibliothek*	Das Hinzuziehen von Lehrwerken oder Besuchen der Bibliothek wird als Strategie benannt.
		SB-QU-GS: *Google-Suche*	Das Suchen mithilfe von Google wird als Strategie benannt.
		SB-QU-YT: *YouTube*	Das Hinzuziehen von YouTube-Videos wird als Strategie benannt.
		SB-QU-IN: *Institutionen*	Das Hinzuziehen von Informationen, die von Institutionen (bspw. WHO) herausgegeben werden, wird als Strategie benannt.
		SB-QU-WP: *Wikipedia*	Das Hinzuziehen von Wikipedia wird als Strategie benannt.
		SB-QU-WS: *Wissensserien*	Das Hinzuziehen von Folgen einer Wissensserie (bspw. Quarks) wird als Strategie benannt.

Auch die Antworten auf die Frage nach den bekannten *naturwissenschaftlichen Arbeits- und Denkweisen* (AD) ermöglichten eine weitere Ausdifferenzierung in Subkategorien. Die Kategorie *Arbeitsweise und Methodik* (AD-AM) umfasst dabei konkrete, handlungsorientierte Aspekte der naturwissenschaftlichen Praxis, also Tätigkeiten, die tatsächlich ausgeführt werden können. Im Gegensatz zu Kategorien, die eher begleitende oder beschreibende Merkmale des Handelns thematisieren, fokussiert die Ka-

tegorie *Arbeitsweise und Methodik* AD-AM auf das aktive Vollziehen naturwissenschaftlicher Methoden und Prozesse. Diese Kategorie ließ sich je nach Schwerpunkt in fünf weitere Facetten unterteilen, während die übrigen acht Kategorien keine zusätzliche Ausdifferenzierung erforderten (Tabelle 34).

Tabelle 34: S1.F3 - Kategorien und Subkategorien des *Verständnisses naturwissenschaftlicher Arbeits- und Denkweisen* (AD)

Überkategorie	Kategorie	Subkategorie	Definition
Verständnis naturwissenschaftlicher Arbeits- und Denkweisen (AD)	AD-AM: *Arbeitsweise und Methodik*	AD-AM-EX: *Experiment*	Das Experiment wird als naturwissenschaftliche Arbeitsweise genannt.
		AD-AM-HY: *Hypothesen*	Das Aufstellen oder Prüfen von Hypothesen wird als naturwissenschaftliche Arbeitsweise genannt.
		AD-AM-PE: *Phänomene erklären*	Das Erkennen und Erklären von Phänomenen wird als naturwissenschaftliche Arbeitsweise genannt.
		AD-AM-PB: *Phänomene beobachten*	Das Beobachten von Phänomenen wird als naturwissenschaftliche Arbeitsweise genannt.
		AD-AM-DE: *Datenerhebung*	Das Erheben und Analysieren von Daten wird als naturwissenschaftliche Arbeitsweise genannt.
	AD-NA: *Naiv/Unkonkret*		Die genannte Vorstellung ist unkonkret oder naiv.
	AD-KO: *Kommunikation/ Diskussion*		Der wissenschaftliche Austausch oder die Kommunikation naturwissenschaftlicher Erkenntnisse werden als naturwissenschaftliche Arbeits- und Denkweise benannt.
	AD-LO: *Logik*		Die Verwendung von Logik wird als naturwissenschaftliche Arbeits- und Denkweise benannt.
	AD-RE: *Reproduzierbarkeit*		Die Reproduzierbarkeit naturwissenschaftlicher Erkenntnisse wird als naturwissenschaftliche Arbeits- und Denkweise benannt.
	AD-VS: *Berücksichtigung verschiedener Sichtweisen/ Interpretationen*		Die Berücksichtigung verschiedener Sichtweisen wird als naturwissenschaftliche Arbeits- und Denkweise benannt.
	AD-OB: *Objektivität*		Objektivität wird als naturwissenschaftliche Arbeits- und Denkweise benannt.
	AD-KP: *Offenheit für Kritik/ Prüfung*		Die Offenheit für Kritik und Prüfung wird als naturwissenschaftliche Arbeits- und Denkweise benannt.
	AD-MO: *Modelle*		Die Verwendung von Modellen wird als naturwissenschaftliche Arbeits- und Denkweise benannt.

Auch die Kategorisierung des *Verständnisses der Begrifflichkeit des Kritischen Denkens* (KD) ergab fünf Hauptkategorien, wobei sich die Kategorie *Hinterfragen* (KD-HF) wiederum in sechs Facetten differenzieren ließ, je nachdem, welche Aspekte konkret hinterfragt wurden (Tabelle 35).

Tabelle 35: S1.F3 - Kategorien und Subkategorien des *Verständnisses Kritischen Denkens* (KD)

Überkategorie	Kategorie	Subkategorie	Definition
Verständnis Kritisches Denken (KD)	KD-HF: *Hinterfragen*	KD-HF-DI: *Diffus*	Es liegt ein diffuses oder nicht formulierbares Verständnis der Begrifflichkeit des Kritischen Denkens vor.
		KD-HF-WI: *Wissen*	Unter der Begrifflichkeit des Kritischen Denkens wird das Hinterfragen von Wissen verstanden.
		KD-HF-ES: *Eigener Standpunkt*	Unter der Begrifflichkeit des Kritischen Denkens wird das Hinterfragen des eigenen Standpunktes verstanden.
		KD-HF-AN: *Ansichten*	Unter der Begrifflichkeit des Kritischen Denkens wird das Hinterfragen von allgemeinen Ansichten verstanden, wobei eigene Ansichten hiervon ausgenommen sind.
		KD-HF-VF: *Verfassende von Informationen*	Unter der Begrifflichkeit des Kritischen Denkens wird das Hinterfragen der Verfassenden von Information oder verbundener Faktoren verstanden.
		KD-HF-AS: *Aussagen*	Unter der Begrifflichkeit des Kritischen Denkens wird das Hinterfragen von Aussagen verstanden.
	KD-SK: *Skepsis/nicht glauben*		Unter der Begrifflichkeit des Kritischen Denkens wird Skepsis im Sinne einer zunächst ungläubigen Grundhaltung verstanden.
	KD-QU: *Quellen*		Unter der Begrifflichkeit des Kritischen Denkens wird das Hinzuziehen von oder das Überprüfen mit Quellen verstanden.
	KD-VS: *Verschiedene Sichtweisen*		Unter der Begrifflichkeit des Kritischen Denkens wird das Berücksichtigen oder Einnehmen mehrerer Sichtweisen verstanden.
	KD-ST: *Änderung des Standpunktes*		Unter der Begrifflichkeit des Kritischen Denkens wird eine offene Haltung gegenüber der Änderung des eigenen Standpunktes verstanden.

Bei der Auswertung wurden zwei verschiedene Zählweisen unterschieden. Zum einen wurde bei den Mehrfachzählungen erfasst, wie häufig eine Kategorie insgesamt kodiert wurde, zum anderen in den Einzelzählungen festgehalten, wie viele Personen eine bestimmte Kategorie erreichten. Bei letzterer Zählweise erhöhten Mehrfachnennungen eines Aspekts einer teilnehmenden Person den zugehörigen Wert nicht.

Diese Methode erwies sich als besonders geeignet, da das Erreichen einer Kategorie darauf hindeutet, dass dieser Aspekt als relevantes Kriterium wahrgenommen wurde. Wiederholungen verdeutlichen zwar die Relevanz für eine Person, beeinflussen jedoch nicht die Aussagekraft auf Gruppenebene. Die Einzelzählungen repräsentieren somit die Anzahl an Teilnehmenden, die einer bestimmten Kategorie zugeordnet werden konnten. Ergänzend zeigt die relative Summe den prozentualen Anteil einer Kategorie an allen erfassten Kodierungen.

Die Antworten auf den CRT wurden analysiert und tabellarisch dargestellt. Dabei wurde die Anzahl richtiger und falscher Antworten erfasst und in Relation zur Gesamtzahl der Teilnehmenden der Stichprobe gesetzt. Als Referenzwerte wurden die in Kapitel 4.2.3 dargestellten Daten herangezogen, die von Frederick (2005) erhoben wurden (s. Tabelle 9).

5.1.3.2 Ergebnisse

Bei der Auswertung der von den Teilnehmenden angegebenen *Kriterien zur Bewertung naturwissenschaftlicher Informationen in Medien* (KB) wurden im Schnitt 3,7 verschiedene Kriterien verzeichnet. Insgesamt waren drei große Schwerpunkte erkennbar. Die *Plattformen* (KB-PL), auf denen Informationen dargestellt werden, waren für 11 von 15 Befragten (73,3 %) ein relevantes Kriterium. Diese Kategorie machte insgesamt 21,5 % aller Kodierungen aus (Tabelle 36). Ein repräsentatives Beispiel dieser Kategorie stellt folgende Aussage dar: „Herausgeber: Ich würde Informationen aus der Bildzeitung bspw. zunächst einmal kritisch hinterfragen, während ich Informationen aus öffentlich-rechtlichen Kanälen höher ansehen würde“ (S1.F3-S22-B3_5-NB16OF, Pos. 5). Insgesamt wurde deutlich, dass allgemein bekannte Verlage und Zeitungen als glaubhafter eingestuft wurden als soziale Medien oder Internetforen, da letztere ohne Prüfung für alle Personen frei zugänglich sind. Für 60 % (9 von 15) der Befragten spielte bei rezipierten Informationen außerdem die *Quellentransparenz* (KB-TR) eine zentrale Rolle. „Für mich ist besonders relevant, ob in schriftlichen Artikeln Quellen angegeben wurden […]“ (S1.F3-S22-B3_5-IP30JE, Pos. 2) dient exemplarisch als Antwort, die dieser Kategorie zugeordnet wurde. Im gleichen Zuge und mit gleicher Verteilung stand auch die *Verlässlichkeit von Informationen oder Quellen* (KB-VL) für die meisten Teilnehmenden im Fokus, wobei Kriterien für diese Glaubwürdigkeit in Aussagen wie „Nach Seriosität der Quelle“ (S1.F3-S22-B3_5-JO16OF, Pos. 3) häufig nicht enthalten waren oder diese eine gewisse Subjektivität implizierten. Der eigene *Wissensstand* (KB-EF) spielte nur bei knapp einem Viertel der Befragten eine Rolle, wohingegen die *Aufmachung und Präsentation von Informationen* (KB-PR)

für fast die Hälfte der Teilnehmenden entscheidend war. Gerade aber die professionelle Aufbereitung ist stark subjektiv geladen und sagt nicht direkt etwas über die inhaltliche Richtigkeit einer Quelle aus. Die *Berücksichtigung mehrerer Standpunkte in Quellen* (KB-VG) oder auch die *Kongruenz mit anderen Quellen* (KB-KG) waren nur für jeweils zwei Personen ein relevantes Kriterium.

Tabelle 36: S1.F3 - Ergebnisse der Fragebogenerhebung, *Kriterien bei der Bewertung naturwissenschaftlicher Informationen in Medien* (KB), *N* = 15

Kategorie/Code	Summe Einfach-zählung	Summe Mehrfach-zählung	Summe rel. %
KB-PL: Plattform/Art der Quelle	11	14	**21,5**
KB-TR: Transparenz/Angabe von Quellen und Daten	9	11	**16,9**
KB-VL: Verlässlichkeit der Information/Quelle	9	10	**15,4**
KB-PR: Professionelle Aufbereitung (Schrift, Form, Präsentation)	7	9	**13,8**
KB-EX: Expertise der*s Autor*in	7	7	**10,8**
KB-EF: Eigenes Fachwissen	4	4	**6,2**
KB-AK: Aktualität	2	3	**4,6**
KB-BK: Bekanntheit der Quelle	2	2	**3,1**
KB-VG: Betrachtung/Vergleich mehrerer Quellen/Standpunkte	2	2	**3,1**
KB-KG: Kongruenz mit weiteren Quellen	2	2	**3,1**
KB-RP: Repräsentativität	1	1	**1,5**
Summe Kriterien	**3,733**	**MW**	

Dennoch fand das *Verwenden von Quellen* (SB-QU) als *Strategie bei der Bewertung naturwissenschaftlicher Informationen* (SB) bei fast allen Teilnehmenden Erwähnung, wenn auch unterschiedlich konkret ausgeprägt. 14 von 15 Proband*innen (93,3 %) führten die *Verwendung von Quellen* (SB-QU) an, wobei der größte Teil, etwa 53 % (8 von 15), *unkonkretisiert* (SB-QU-UK) ließ, welche Quellen konkret Anwendung finden (Tabelle 37). Eine unkonkrete Aussage stellt das Beispiel „Ich suche mir seriöse Quellen raus und lese mir diese ausführlich durch“ (S1.F3-S22-B3_5-TQ32KF, Pos. 4) dar. Als konkrete Quellen wurden hingegen *Forschungsliteratur* (SB-QU-FL), wie wissenschaftliche Paper, oder *Lehrwerke* (SB-QU-LB) von jeweils vier Befragten genannt.

Die *Nutzung verschiedener Quellen und Standpunkte* (SB-VQ) fand bei drei Personen explizit Erwähnung.

Abseits einer selbstständigen Quellenrecherche stand der „Austausch mit Freunden, Familie, Kollegen“ (S1.F3-S22-B3_5-IF41SB, Pos. 5) an erster Stelle, was unter die Kategorie *Austausch mit anderen (Laien)* (SB-AT-LA) fiel, da sich dieser Austausch von der „Absprache mit Experten bspw. dem eigenen Arzt/Ärztin“ (S1.F3-S22-B3_5-JO16OF, Pos. 11) hinsichtlich der Expertise unterschied. 40 % (6 von 15) suchten dementsprechend Rat in ihrem sozialen Umfeld (SB-AT-LA) und nur 20 % (3 von 15) suchten den Austausch mit *Expert*innen* (SB-AT-EX).

Tabelle 37: S1.F3 - Ergebnisse der Fragebogenerhebung, *Strategien bei der Bewertung naturwissenschaftlicher Informationen in Medien* (SB), *N* = 15

Kategorie/Code		Summe Einfach-zählung	Summe Mehrfach-zählung	Summe rel. %	
SB-VQ: Nutzung verschiedener		3	3	**6,8**	
SB-OF: Offenheit gegenüber neuen		2	2	**4,5**	
SB-ÜW: Überprüfung des eigenen Wissenstandes		2	2	**4,5**	
SB-AT: Austausch	SB-AT-LA: Austausch mit anderen (Laien)	6	6	**13,6**	**20,5**
	SB-AT-EX: Austausch mit anderen (Expert*innen)	3	3	**6,8**	
SB-QU: Quellen	SB-QU-UK: Unkonkret	8	9	**20,5**	**63,6**
	SB-QU-FL: Forschungsliteratur	4	5	**11,4**	
	SB-QU-LB: Lehrwerke/ Bibliothek	4	4	**9,1**	
	SB-QU-GS: Google-Suche	3	3	**6,8**	
	SB-QU-YT: YouTube	2	2	**4,5**	
	SB-QU-SM: Soziale Medien	2	2	**4,5**	
	SB-QU-IN: Institutionen (bsp. WHO)	1	1	**2,3**	
	SB-QU-WP: Wikipedia	1	1	**2,3**	
	SB-QU-WS: Wissensseiten (z.B. Quarks)	1	1	**2,3**	
Summe Strategien		**2,800**	**MW**		

Bei der Untersuchung des *Verständnisses naturwissenschaftlicher Arbeits- und Denkweisen* (AD) zeichnete sich ein klares Bild, insbesondere in Bezug auf die assoziierten *Arbeitsweisen und Methoden* (AD-AM) (Tabelle 38). Das *Experiment* (AD-AM-EX) wurde von 80 % (12 von 15) der Befragten als elementarer und kennzeichnender Bestandteil naturwissenschaftlichen Arbeitens verstanden. Eine Antwort lautete beispielsweise: „Das Ausführen von Versuchen im Labor ist mir als naturwissenschaftliche Arbeitsweise bekannt“ (S1.F3-S22-B3_5-TQ32KF, Pos. 5). Etwa 47 % (7 von 15) nannten zudem das *Aufstellen und Prüfen von Hypothesen* (AD-AM-HY) als exemplarische Arbeitsweise. Das *Erkennen* (AD-AM-PE) und *Beobachten von Phänomenen* (AD-AM-PB) wurde von je drei Personen angeführt.

„Wissenschaftlicher Austausch“ (S1.F3-S22-B3_5-HP20OF, Pos. 12) wurde unter der Kategorie *Diskussion und Kommunikation* (AD-KO) verbucht und von drei Teilnehmenden als naturwissenschaftliches Kernelement verstanden.

Tabelle 38: S1.F3 - Ergebnisse der Fragebogenerhebung, *Verständnis naturwissenschaftlicher Arbeits- und Denkweisen* (AD), *N* = 15

Kategorie/Code		Summe Einfachzählung	Summe Mehrfachzählung	Summe rel. %	
AD-AM: Arbeitsweise und Methodik	AD-AM-EX: Experiment	12	12	**24,5**	**53,1**
	AD-AM-HY: Aufstellen/Prüfen von Hypothesen	7	7	**14,3**	
	AD-AM-PE: Erkennen und erklären von Phänomenen	3	3	**6,1**	
	AD-AM-PB: Beobachten von Phänomenen	3	3	**6,1**	
	AD-AM-DE: Datenerhebung-/analyse	1	1	**2,0**	
AD-NA: Naiv/Unkonkret		7	7	**14,3**	
AD-KO: Diskussion/Kommunikation		3	4	**8,2**	
AD-LO: Logik		2	3	**6,1**	
AD-RE: Reproduzierbarkeit		3	3	**6,1**	
AD-VS: Berücksichtigung verschiedener Sichtweisen/Interpretationen		2	2	**4,1**	
AD-OB: Objektivität		2	2	**4,1**	
AD-KP: Offenheit für Kritik/Prüfung		1	1	**2,0**	
AD-MO: Verwendung von Modellen		1	1	**2,0**	

Rund 47 % (7 von 15) der Teilnehmenden tätigten mit 14,3 % der Gesamtkodierungen *naive oder unkonkrete Aussagen* (AD-NA) zu naturwissenschaftlichen Arbeits- und

Denkweisen, wie beispielsweise die Benennung von „Forschung im labor [sic]“ (S1.F3-S22-B3_5-HP20OF, Pos. 13) als Inbegriff wissenschaftlicher Arbeit oder die folgende Aussage zeigt:

> „Forschen an sich ist eine "wissenschaftliche" Arbeitsweise. Man geht an ein Thema dran, dass [sic] man versucht ohne vorher erwartete Ergebnisse zu bearbeiten. Dabei werden möglichst viele Einflussfaktoren betrachtet [sic] die man aber (vor allem in der Physik) einfach danach raus streicht [sic] wenn sie einem nicht passen. (nicht relevante Messergebnisse)“ (S1.F3-S22-B3_5-NB27LF, Pos. 13).

Diese ist in mehrerlei Hinsicht problematisch, da sich in dieser Aussage ein verzerrtes Bild wissenschaftlichen Arbeitens abzeichnet und eine Willkürlichkeit unterstellt wird.

Zusammengenommen wird ein recht einseitiges Bild naturwissenschaftlichen Arbeitens und Denkens gezeichnet, das sich primär auf Methoden und Arbeitsweisen fokussiert. Erweiterte epistemologische Perspektiven, wie sie im Sinne von Nature of Science (NOS) diskutiert werden (vgl. Kap. 2.1.2), finden, wie schon in anderen Studien gezeigt (Abd-El-Khalick, 2006; Liang et al., 2009; McComas, 1998b; S. Müller, 2021), kaum Erwähnung. Insbesondere menschliche, dynamische und gesellschaftlich geprägte Facetten, wie die Vorläufigkeit, die soziokulturelle Eingebundenheit oder die Kreativität, finden keinerlei Erwähnung.

Das *Verständnis von Kritischem Denken* (KD) zeigte sich bei den Proband*innen als sehr inkonsistent und heterogen, da sehr unterschiedliche Vorstellungen vertreten waren. Am häufigsten wurde das *Hinterfragen* (KD-HF) von etwa 73 % (11 von 15) der Befragten als Synonym für Kritisches Denken interpretiert, wobei dieser Begriff häufig *diffus* (KD-HF-DI) und ohne Bezugsrahmen verwendet wurde (6 von 15 Befragten) (Tabelle 39). Exemplarisch kann hier das folgende Zitat angeführt werden: „Kritisch macht es [sic] wenn man Dinge neu hinterfragt und nicht einfach stumpft [sic] anderen Menschen/ Meinungen hinterher rennt [sic]. Aber auch noch realistisch sein und nicht jedem Kontra hinterherrennen“ (S1.F3-S22-B3_5-BO38HB, Pos. 17). Nur zwei Teilnehmende bezogen sich konkret auf *Wissen* (KD-HF-WI), zwei weitere auf das *Hinterfragen des eigenen Standpunktes* (KD-HF-ES) und wiederum zwei auf das *Hinterfragen der Verfassenden von Informationen* (KD-HF-VF).

Tabelle 39: S1.F3 - Ergebnisse der Fragebogenerhebung, *Verständnis Kritisches Denken*, *N* = 15

Kategorie/Code		Summe Einfach-zählung	Summe Mehrfach-zählung	Summe rel. %	
KD-HF: Hinterfragen (HF)	KD-HF-DI: Hinterfragen diffus	6	7	**23,3**	**50,0**
	KD-HF-WI: Wissen hinterfragen	2	2	**6,7**	
	KD.HF-ES: Eigenen Standpunkt hinterfragen	2	2	**6,7**	
	KD-HF-AN: Ansichten hinterfragen	1	1	**3,3**	
	KD-HF-VF: Verfasser von Informationen und verbundene Faktoren hinterfragen	2	2	**6,7**	
	KD-HF-AS: Aussagen hinterfragen	1	1	**3,3**	
KD-SK: Skepsis/nicht alles glauben		5	5	**16,7**	
KD-QU: Quellen hinzuziehen		5	5	**16,7**	
KD-VS: Einbezug verschiedener Sichtweisen		4	4	**13,3**	
KD-ST: Offenheit für Änderungen des Standpunktes		1	1	**3,3**	

Kritisches Denken wurde zudem von einem Drittel der Befragten als *Skepsis* (KD-SK) interpretiert, wie das Zitat „Wer kritisch denkt, glaubt nicht alles, was er [sic] oder ihr vorgesetzt wird [sic] sondern schaut über den Tellerrand hinaus“ (S1.F3-S22-B3_5-NB16OF, Pos. 15) zeigt. Auch diese Skepsis zeigte sich deutlich ungerichtet und kann eher als eine ungläubige Haltung interpretiert werden. Für ein weiteres Drittel der Proband*innen bedeutet Kritisches Denken „vielmehr selbst nachzuforschen und eventuell in anderen Quellen das Thema nachzulesen“ (S1.F3-S22-B3_5-TQ32KF, Pos. 15), was unter die Kategorie *Quellen hinzuziehen* (KD-QU) fiel.

Im Gesamten kann festgestellt werden, dass die teilnehmenden Studierenden keine einheitliche Vorstellung von der Begrifflichkeit des Kritischen Denkens aufwiesen. Dieser Umstand deutet auf deutlich subjektiv geladene, individuelle Definitionen der*s Einzelnen hin, was das Wahrnehmen der Vermittlerrolle im Kontext des Kritischen Denkens erschwert.

Der Einsatz des bereits etablierten *Cognitive Reflection Tests* (CRT) von Frederick (2005) zeigte eine annähernd gleichmäßige Verteilung der Ergebnisse. Tabelle 40 gibt

einen Überblick über die korrekt beantworteten Fragen pro Fragestellung. Insgesamt wurden von den Teilnehmenden 44,4 % der Fragen richtig beantwortet, was einem Durchschnitt von 1,3 korrekten Antworten pro Person entspricht. Ein Drittel der Studierenden konnte keine der drei CRT-Fragen korrekt beantworten und wählte die intuitive Antwort, während ein Fünftel alle Fragen richtig löste. Bei etwa 50 % der Befragten (keine oder nur eine richtige Antwort) deutet dies auf eine eingeschränkte Erstimpulskontrolle beim Lösen der Aufgaben hin. Die inkorrekt beantworteten Antworten entsprachen ausnahmslos dem erwarteten Muster der intuitiven Erstantworten, weshalb nicht von Rechenfehlern seitens der Teilnehmenden ausgegangen werden kann. Diese Ergebnisse sind vergleichbar mit den an anderen Universitäten erhobenen Daten von Frederick (2005), die hier als Referenz herangezogen wurden.

Tabelle 40: S1.F3 - Ergebnisse der Fragebogenerhebung, CRT, *N* = 15

Code	CRT1	CRT2	CRT3	Korrekt	%
Korrekt	5	9	6	20	**44,4**

Verteilung der Ergebisse in %			
33,3	20,0	26,7	20,0
0	1	2	3
Low			High

5.1.4 Vermittlungseinheit

Im Seminar zu fachbezogenen Lern- und Kommunikationsprozessen im Sommersemester wurde neben den Erhebungen die Vermittlungseinheit erstmalig erprobt (Tabelle 41)(vgl. Kap. 4.2.5). Diese beinhaltete noch keine Methode der Systemisch-reflexiven Stellungnahme, sondern stattdessen ein Rollenspiel, das mit einer Web-Quest kombiniert wurde. Für die Konzeption des Rollenspiels diente der Methodenpool der Universität zu Köln als Orientierung (Reich, 2017). Im Zentrum dieses Rollenspiels stand die Fragestellung, ob der Handel und die Förderung Seltener Erden erweitert werden sollten, zu deren Beantwortung die Studierenden in vier Gruppen eingeteilt wurden. Diesen Gruppen wurden randomisiert Steckbriefe für Positionen ausgeteilt, die diese im weiteren Verlauf einnehmen sollten. Darunter waren zwei Pro-Positionen, verkörpert durch eine Ingenieurin und einen Firmenchef, und zwei Kontra-Positionen, welche durch eine Politikerin der Grünen und einen in einem Abbaugebiet lebenden chinesischen Bauern repräsentiert wurden. Die Aufgabenstellungen lauteten wie folgt:

1. **Lest** euch den ausgeteilten Steckbrief in Ruhe durch und versucht, euch in diese Person **hineinzuversetzen**.
   Was bewegt die Person? Wie würde sie argumentieren?
2. Geht ins Internet und **sammelt Informationen** und Argumente, die eure Position stützen. Bereitet euch auf eine Podiumsdiskussion vor, in der ihr eure **Interessen vertreten** müsst!
   Was sind eure Argumente? Wie könnt ihr eure Position verteidigen?
3. **Sammelt eure Argumente**/Notizen/Zusatzwissen schriftlich in einem Word-Dokument und **ladet** es vor der Podiumsdiskussion **hoch.**
   (Name: „GRUPPE NR_IDENTIFIKATION_NUMMER"
   Beispiel: GRUPPE 1_PRO_1)

Jede Gruppe arbeitete bei der Recherche zusammen und bestimmte am Ende eine repräsentierende Person, die die Diskussion stellvertretend führte, und eine protokollierende Person, die den Diskussionsablauf stichpunktartig verfolgte. Die Diskussion des Rollenspiels fand dann mit den Repräsentant*innen der vier Positionen unter der genannten zentralen Fragestellung statt. Während der Diskussion im Rollenspiel gab es für die Teammitglieder einen Hot-Seat, auf dem die nicht aktiven Mitglieder jederzeit Platz nehmen und ihr Argument in die Diskussion einbringen konnten.

Nach der Diskussionsrunde wurden die Teilnehmenden aus ihrer Position entlassen, um eine Distanz zur vorher eingenommenen Rolle einnehmen zu können und die Kritik am Rollenverhalten von der eigenen Person zu trennen (Reich, 2017). Im Anschluss fand eine Reflexionsphase statt, in der die gemachten Erfahrungen und Beobachtungen aus dem Rollenspiel rekapituliert werden sollten.

Das Ziel bestand darin, den Teilnehmenden durch Selbsterfahrung aufzuzeigen, dass auch Recherchen und Diskussionen, die auf der Grundlage zuvor eingeholter Informationen und vermeintlicher Fakten geführt werden, standpunktabhängigen Verzerrungstendenzen unterliegen können. Das Rollenspiel diente somit als bewusst inszeniertes Beispiel für das Auftreten eines Confirmation Bias, um diesen erfahrbar zu machen.

Tabelle 41: S1 - Übersicht über die Seminarsitzungen im SoSe22, Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP)

Sitzung	Inhalt
1	Begrüßung, Organisatorisches und Einführung – Fragebogen S1.F1. – Einblick in die Kompetenzbereiche und Ziele der Chemiedidaktik
2	Unterrichtsmethoden
3	– Individuelle Förderung: Differenzierung und Diagnostik im (inklusiven) Chemieunterricht – Fragebogen S1.F2
4	– Umgang mit Medien im Chemieunterricht: Experimente, Modelle, didaktische Programme und mehr – Fragebogen S1.F3
5	Nature of Science 1 – Workshop; angeleitete Selbsterfahrung mit NOS
6	Nature of Science 2 – Diskussion und Reflexion von NOS
7	Science and Pseudoscience – Das Abgrenzungsproblem und Annäherungen an Differenzierungsmöglichkeiten
8	Rollenspiel
9	News-Test – Einstieg kognitive Verzerrungen
10	Kognitive Verzerrungen – Was unser Denken und Verhalten beeinflusst; Dual-Process Theory
11	Kritisches Denken – Begriffsdefinition und Relevanz bei der Bewertung naturwissenschaftlicher Kontexte
12	Weiterführung zum Kritischen Denken und Rückblick auf das Seminar
13	– Reflexion des Seminars – Evaluierung, Ausblick und Abschluss

Das Rollenspiel erzielte die gewünschte Wirkung und die Studierenden gaben in Wortbeiträgen an, dass sie im Rechercheprozess, aufgrund des hohen Fokus auf die einzunehmende Position, alternative Standpunkte ausblendeten. Allerdings fiel es den Gruppen schwer, Positionen abseits der eigenen Überzeugung einzunehmen und diese im Diskussionsverlauf überzeugend zu vertreten. Trotz seines Potenzials als effektives Mittel zur Selbsterfahrung erwies sich das Rollenspiel als sehr zeitintensiv, weshalb es in nachfolgenden Vermittlungseinheiten durch alternative Methoden ersetzt wurde.

5.1.5 Abschluss der ersten Studie

Auf Grundlage der bisherigen Ergebnisse lässt sich feststellen, dass in Bewertungssituationen mit naturwissenschaftlichem Bezug ohne explizite Thematisierung subjektive Bewertungsansätze dominieren. Eigenes naturwissenschaftliches Wissen wird selten herangezogen und reflexive Fragestellungen zeigen sich größtenteils ich-bezogen. Dieser Umstand lässt sich jedoch nicht auf die fehlende Fähigkeit zurückführen, naturwissenschaftlich ausgerichtete Fragen zu stellen, sondern scheint entweder dem Nichterkennen oder Nichtbeachten naturwissenschaftlicher Bezüge geschuldet zu sein. Gleichzeitig lieferte die Auswertung des CRT Hinweise auf eine geringe Erstimpulsinhibition bei etwa der Hälfte der Teilnehmenden, was im Hinblick auf die gemachten Ergebnisse den Schluss nahelegt, dass den affektiven Erstimpulsen bei nicht explizit thematisierter Kontextualisierung nachgegeben wird und trotz der grundsätzlichen Fähigkeit kein naturwissenschaftliches Fachwissen in den Bewertungsprozess einfließt. In der Realität außerhalb des Klassenraums werden jedoch selten Impulse und Richtungen bei Bewertungsprozessen vorgegeben, weswegen sich weiter die Frage stellt, wie das vorhandene Wissen tatsächlich in Entscheidungssituationen zum Tragen kommen kann.

Bei der Auseinandersetzung mit naturwissenschaftlichen Informationen scheinen vornehmlich quellenbezogene Kriterien und Strategien ausschlaggebend zu sein, wie die Plattform der Publikation, die Transparenz oder die Darstellung einer Ressource. Neben dem Hinzuziehen von Quellen stellt aber auch der Austausch mit Personen des sozialen Umfelds eine gängige Strategie dar, die sich in Übereinstimmung mit den Erkenntnissen von Ajzen (1991) und Elster (1989, 2008) (vgl. Kap. 2.3) als erwartbares Verhalten zeigt.

Das Verständnis naturwissenschaftlicher Arbeits- und Denkweisen kann als grundlegend methodisch ausgerichtet und prozessorientiert beschrieben werden, wobei tiefere menschliche und gesellschaftliche Dimensionen naturwissenschaftlichen Arbeitens unbeachtet blieben. Da dieser Aspekt bereits in mehreren Studien Gegenstand der Untersuchungen war (bspw. Abd-El-Khalick, 2006; Liang et al., 2009; McComas, 1998b; S. Müller, 2021), wird in den folgenden Erhebungen zugunsten einer stärkeren Fokussierung auf andere Bereiche auf eine weitere Betrachtung verzichtet.

Das Verständnis Kritischen Denkens ist ebenfalls eher als diffus und unkonkret einzustufen, wobei eine grundlegende Skepsis häufig als zentraler Bestandteil Kritischen Denkens verstanden wird.

Mit Blick auf die Methodik der Fragengenese stellt sich die Formulierung von Fragen als vorteilhaft dar, da auf diese Weise kein Wissen aktiviert werden muss, welches zur Bewertung des Kontextes eventuell nicht vorhanden ist und so einen limitierenden Faktor darstellen würde. Stattdessen kann vielmehr die Denkrichtung der Proband*innen ergründet sowie ein Einblick in die Fragen, die diese sich selbst im Bewertungsprozess stellen würden, gewonnen werden. Diese Unabhängigkeit vom ansonsten notwendigen Verfügungswissen stellt in der weiteren Erhebung und Erprobung in anderen Kontexten einen großen Vorteil dar. Das Format des Fragebogens ist jedoch sehr stark limitiert und bietet keine Möglichkeiten zur Vertiefung oder externen Nachfrage, weshalb im weiteren Verlauf von Studie 2 ein Interviewformat erprobt wurde.

Die grundsätzliche Konzeption der Vermittlungseinheit erwies sich im Seminarverlauf als vielversprechend, wobei auf Grundlage der gemachten Erfahrungen Anpassungen an den Aufgabenstellungen und der zeitlichen Verteilung im Vermittlungsprozess vorgenommen wurden. Zudem führten die Erkenntnisse aus dem Rollenspiel zu weiteren methodischen Anpassungen, wie der Verwerfung der Methode zugunsten der Erprobung neuer Ansätze.

5.2 Studie S2

Abbildung 19: Übersicht über den Aufbau und die Erhebungsparameter der Studie S2

Die Studie S2 verkörperte im Wintersemester 2022/2023 eine Neuausrichtung der Erhebungsinstrumente nach der initialen Studie S1. Sie bestand aus einer reduzierten Fragebogenbefragung, in welche neben dem bereits zur Anwendung gekommenen CRT von Frederick (2005) auch erstmalig der SRIS von Silvia (2021) einfloss. Da es sich um bereits etablierte Erhebungsinstrumente handelte und keine eigenen Elemente in den Fragebogen der Studie S2 integriert wurden, war eine Pilotierung des Fragebogens nicht erforderlich. Neben der Erhebung durch das Fragebogenformat

wurde erstmalig eine Interviewstudie durchgeführt. Die hierfür entwickelten Interviewleitfäden wurden zunächst mit $N = 3$ Studierenden aus der Vorlesung *Mathematisch-Naturwissenschaftliche Grundlegung (MNF-B)* pilotiert. Anschließend wurden die Interviews sowohl mit Bachelorstudierenden aus dem Seminar *Grundlegende Aspekte der Fachdidaktik (B2 HR/SP B3 GG)* als auch Masterstudierenden aus dem Seminar *Ausgewählte Themengebiete des Chemieunterrichts (M1 HR/GG)* durchgeführt (vgl. Kap. 4.2.2). Beide Seminare wurden von Dozierenden der Chemiedidaktik geleitet und der Untersuchende trat nicht als Lehrperson in Erscheinung. Die Gesamtstichprobe betrug $N = 10$, wobei $n = 4$ Studierende aus dem Masterseminar und $n = 6$ Studierende aus dem Bachelorseminar stammten. Alle zehn Proband*innen füllten ebenfalls den in S2 verwendeten Fragebogen aus und nahmen an der Interviewstudie teil (Abbildung 19).

Auch diese Studie kann noch als Vorstudie angesehen werden, da hier keine Definitionen als Grundlage Anwendung fanden und die Untersuchung rein induktiv vorgenommen wurde. Dennoch erfolgt eine detailliertere Betrachtung der Studie, da sie die Vorstufe zur Entwicklung der ersten Hauptstudie S3 darstellt und erste Vorläufer der späteren Kategorien erkennen lässt.

5.2.1 Interviewerhebung S2.I1

Im Folgenden wird ein Überblick über die durchgeführte Interviewstudie S2.I1 gegeben. Eine detaillierte Darstellung des Interviewaufbaus findet sich dazu bereits in Kapitel 4.2.4.. Die in der Untersuchung verwendeten Interviewkontexte (vgl. Kap. 4.2.6) umfassten:

Kontext A1: Solarenergieversorgung vs. Windkraftanlagen

Kontext A2: Solarenergieversorgung vs. Windkraftanlagen vs. Atomenergie vs. Kohleverbrennung

Kontext B: E-Mobilität vs. Brennstoffzellenantrieb

Die Erweiterung des Kontextes A1 um A2, welche nur in dieser Vorstudie vorgenommen wurde, strebte die Untersuchung an, ob die gesellschaftlich stark emotional aufgeladenen Positionierungen gegenüber Atomenergie und Kohleverbrennung eine unbewusste Beeinflussung hervorrufen und dadurch eine stärkere Vorprägung der Antworten bewirken, die wiederum zu weniger reflektierten Entscheidungen führen könnten (vgl. Kap. 4.2.6.2). Die Kontexte wurden den Teilnehmenden randomisiert zugeteilt.

Mit Rückbezug auf die in Studie 1 gewonnenen ersten Einblicke und rekurrierend auf theoretische Vorüberlegungen im Kontext der Dual-Process Theory sowie des Phänomens der Knowledge-Action Gap kommt der Selbstreflexion als verbindendem Element zwischen affektiver und kognitiver Ebene eine besondere Bedeutung zu (vgl. Kap. 3.3 und 3.5). Aus diesem Grund wurde in der hier durchgeführten Interviewerhebung der Untersuchung selbstreflexiver Aspekte in Bewertungskontexten ein übergeordnetes Erkenntnisinteresse beigemessen.

Zusätzlich umfasste die Untersuchung das Verständnis der Begrifflichkeiten *Selbstreflexion* sowie der Phrase *kritisch sein*. Letztere wurde im Rahmen des Interviewsettings von der Erhebung des Verständnisses des Konstruktbegriffs hin zur Erfassung des Verständnisses der adjektivisch verwendeten Phrase verlagert. Dieser Ansatz basiert auf der Beobachtung, dass in den Bildungsstandards und Kernlehrplänen mehrerer Bundesländer (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2019, 2022; Sächsisches Staatsministerium für Kultus, 2022; Sekretariat der Kultusministerkonferenz, 2019; Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2024b, 2020) Kritisches Denken nicht explizit als Konstrukt benannt, sondern über adjektivische Formulierungen wie *kritisch sein* oder *kritisch bewerten* implizit adressiert wird. Um eine größtmögliche Nähe zu den im Bildungsdiskurs und alltäglichen Sprachgebrauch etablierten Formulierungen herzustellen, wurde daher gezielt nach dem Verständnis dieser verbreiteten Phrase gefragt.

Ferner wurde eine Selbsteinschätzung der Befragten hinsichtlich ihrer Bewertungsfähigkeit sowie ihres Fachwissens im jeweiligen Kontext erhoben, um potenzielle Einflüsse heuristischer Bewertungsstrategien sichtbar zu machen. Wenn eine Person angibt, sich in der Lage zu fühlen, eine Bewertung vorzunehmen, jedoch im Verlauf des Interviews mangelndes oder kein Fachwissen zum Thema einräumt, legt das den Schluss nahe, dass die Bewertung nicht auf einer fachlichen, sondern auf einer heuristischen Ebene erfolgt.

Insgesamt sollten folgende Untersuchungsfragen durch die Interviewerhebung beantwortet werden:

1. Inwieweit findet bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug Selbstreflexion statt?

2. Welches Verständnis haben Chemielehramtsstudierende von der Begrifflichkeit *kritisch sein*?

3. Welches Verständnis haben Chemielehramtsstudierende vom Begriff der *Selbstreflexion*?

4. Wie schätzen Studierende des Chemielehramtes ihr Fachwissen in dem gestellten Setting ein?

5. Wie schätzen Studierende des Chemielehramtes ihre Bewertungsfähigkeit in Bezug auf den zu bewertenden Kontext ein?

5.2.1.1 Auswertung

Zu Beantwortung der Untersuchungsfrage 1

> Inwieweit findet bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug Selbstreflexion statt?

wurden die nach der Durchführung des Interviews erstellten Transkripte mittels der induktiven Kategorienbildung nach Kuckartz (2016) analysiert vgl. Kap. 4.3.1). Obwohl aus der Voruntersuchung in Studie S1 bereits Kategorien vorlagen, wurde infolge des veränderten Erhebungsformats, welches zu einer abweichenden Datenbasis führte, ein induktiver Ansatz gewählt, um der Zielsetzung nachzukommen, das neue Material ohne Vorannahmen und mit größtmöglicher Offenheit zu analysieren.

Bei der Kodierung wurden zwei übergeordnete Kategorien definiert (Tabelle 42). Die Kategorie der *introspektiven Überprüfung* (SR-IU) fasst alle Aussagen zusammen, die auf eine Einsichtnahme in das persönliche Innenleben hindeuteten. Diese reflexiven Anteile wurden wiederum in zehn Facetten unterschieden und als Subkategorien definiert. Diese subjektiv ich-bezogenen Anteile sind direkte Indikatoren für ein spezifisches selbstreflexives Vorgehen und zeigen so nicht nur auf, ob Selbstreflexion stattfand, sondern auch, welche Ausrichtung diese hatte. Exemplarisch können hier die Subkategorien der *emotionalen Einsicht* (SR-IU-EE), der *eigenen Erfahrungen* (SR-IU-EF) oder der eigenen *Moral und Werte* (SR-IU-MW) angeführt werden, die die Einsichtnahme in die jeweiligen Aspekte umfassen. Die zweite übergeordnete Kategorie

hingegen ist das *Bewusstsein für externe und interne Einflussfaktoren* (SR-EF). Diese Kategorie umspannt Äußerungen, die nahelegten, dass sich die Befragten bestimmter Einflussfaktoren im Bewertungsprozess bewusst waren. Auch diese Kategorie wurde weiter spezifiziert und umfasst elf differenzierte Subkategorien, wobei eine jede für einen spezifischen reflektierten Einflussfaktor steht, beispielsweise die *soziale Norm* (SR-EF-SN), die *Medien* (SR-EF-ME) oder auch die *persönliche Befangenheit* (SR-EF-PB).

Tabelle 42: S2.I1 - Kategorien und Subkategorien der *Selbstreflexion bei der Bewertung* gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (SR)

Überkategorie	Kategorie	Subkategorie	Definition
Selbstreflexion bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (SR)	SR-IU: *Introspektive Überprüfung*	SR-IU-FW: *Bewertung des eigenen Fachwissens*	Das eigene Fachwissen wird bei der Bewertung reflektiert.
		SR-IU-PR: *Persönliche Relevanz (Vor- und Nachteile)*	Die Relevanz für sich selbst im Sinne von Vor- und Nachteilen wird bei der Bewertung reflektiert.
		SR-IU-MW: *Moral und Werte*	Die eigenen Werte und Moralvorstellungen werden bei der Bewertung reflektiert.
		SR-IU-RH: *Retrospektive Handlungsbetrachtung*	Die eigenen Handlungen werden bei der Bewertung rückwirkend betrachtet.
		SR-IU-EE: *Emotionale Einsicht*	Die eigene Gefühlswelt wird bei der Bewertung reflektiert.
		SR-IU-DG: *Diffuses Gefühl*	Ein diffuses Gefühl wird bei der Bewertung als Begründung angegeben.
		SR-IU-EF: *Eigene Erfahrungen*	Die eigenen gemachten Erfahrungen werden bei der Bewertung reflektiert.
		SR-IU-ID: *Identifikation*	Das Selbstbild oder die Möglichkeit, sich mit einem Standpunkt zu identifizieren, werden bei der Bewertung reflektiert.
		SR-IU-IN: *Interesse*	Das eigene Interesse wird bei der Bewertung reflektiert.
		SR-IU-EN: *Eigene Notwendigkeit*	Die eigene Notwendigkeit der Bewertung oder Handlung wird bei der Bewertung reflektiert.
	SR-EF: *Bewusstsein für externe und interne Einflussfaktoren*	SR-EF-SN: *Soziales Umfeld/ Soziale Norm*	Es wird bei der Bewertung ein Bewusstsein für den Einfluss des sozialen Umfelds erkennbar.
		SR-EF-ME: *Medien*	Es wird bei der Bewertung ein Bewusstsein für den Einfluss durch Medien erkennbar.
		SR-EF-PB: *Persönliche Befangenheit/Vorurteile*	Es wird bei der Bewertung ein Bewusstsein für den Einfluss persönlicher Befangenheiten, wie etwa Vorurteilen, erkennbar.

		SR-EF-IN: *Interesse*	Es wird bei der Bewertung ein Bewusstsein für den Einfluss des eigenen Interesses erkennbar.
		SR-EF-ER: *Kindheit und Erziehung*	Es wird bei der Bewertung ein Bewusstsein für die Einflüsse von Kindheit und Erziehung erkennbar.
		SR-EF-SY: *Sympathie*	Es wird bei der Bewertung ein Bewusstsein für den Einfluss von Sympathien erkennbar.
		SR-EF-BI: *Bildung*	Es wird bei der Bewertung ein Bewusstsein für den Einfluss der eigenen Bildung erkennbar.
		SR-EF-PM: *Persönlicher Möglichkeitsrahmen*	Es wird bei der Bewertung ein Bewusstsein für den Einfluss des eigenen Möglichkeitsrahmens erkennbar.
		SR-EF-EX: *Expert*innen-meinungen*	Es wird bei der Bewertung ein Bewusstsein für den Einfluss von Expert*innenmeinungen erkennbar.
		SR-EF-EF: *Erfahrungen*	Es wird bei der Bewertung ein Bewusstsein für die Einflüsse von gemachten Erfahrungen erkennbar.
		MW: *Moral und Werte*	Es wird bei der Bewertung ein Bewusstsein für den Einfluss der eigenen Moral- und Wertvorstellungen erkennbar.

Die Differenzierung in die *introspektive Überprüfung* (SR-IU) und das *Bewusstsein für externe und interne Einflussfaktoren* (SR-EF) wurde vorgenommen, da reflexive Gedankengänge auch ohne expliziten Ausdruck eines Bewusstseins für beeinflussende Faktoren stattfinden können. So verdeutlichen die formulierten Fragen "[W]ie habe ich mich da gefühlt? Habe ich mich wohlgefühlt? Habe ich mich unwohl gefühlt?" (S2.I1-S22-B3_5- BN10LB, Pos. 66) zwar, dass eine *emotionale Einsichtnahme* (SR-IU-EE) stattgefunden hat, jedoch wird nicht explizit hervorgehoben, dass die Emotionen auch einen Einfluss auf die Entscheidungen nehmen können. Im Gegensatz dazu zeigt folgendes Zitat deutlich ein explizites Bewusstsein für externe Einflussfaktoren: „Genauso würde ich aber auch im sozialen Umfeld sagen klar, wir lassen uns irgendwie immer unbewusst beeinflussen, auch wenn man es vielleicht gar nicht möchte.“ (S2.I1-S22-B3_5-XF21FM, Pos. 86). Dieses Beispiel beschreibt explizit den Einfluss des *sozialen Umfelds* (SR-EF-SN) und spiegelt ein bewusstes Wahrnehmen dieser Dynamik wider. Andererseits bedeutet ein klares Benennen von derlei Einflussfaktoren nicht zwangsweise, dass innerhalb des Interviews auch eine derartige Reflexion vorgenommen wurde, weshalb eine Differenzierung beider Ansätze sinnvoll erschien.

Zur Untersuchungsfrage 2

> Welches Verständnis haben Chemielehramtsstudierende von der Begrifflichkeit kritisch sein?

wurden sechs verschiedene Kategorien gebildet (Tabelle 43). Die im Fragebogen vorgenommene Differenzierung der Kategorie *Hinterfragen* (VK-HF) hinsichtlich der hinterfragten Medien wurde im Interviewkontext nicht beibehalten, da diese Unterscheidung keinen inhaltlichen Mehrwert lieferte. Sie fokussierte ausschließlich auf die hinterfragten Medien und nicht auf das Vorgehen des Hinterfragens im Kontext Kritischen Denkens selbst, weshalb die Kategorie in der Interviewerhebung undifferenziert belassen wurde. Weitere Kategorien waren unter anderem das Verständnis der Begrifflichkeit *kritisch sein* als *ungläubige Haltung* (VK-NG) oder auf sich selbst bezogen als *Selbstkritik* (VK-SK). Antworten, die erkennen ließen, dass keine Definitionsvorstellung vorlag, wurden der Kategorie *Nicht definierbar* (VK-ND) zugeordnet.

Tabelle 43: S2.I1 - Kategorien des *Verständnisses von kritisch sein* (VK)

Überkategorie	Kategorie	Definition
Verständnis kritisch sein (VK)	VK-HF: *Hinterfragen*	„Kritisch sein" wird als das Hinterfragen von Dingen oder Aussagen verstanden.
	VK-WW: *Eigenen Werten widersprechend*	„Kritisch sein" wird als eine Bewertung verstanden, die den eigenen Werten widerspricht.
	VK-SK: *Selbstkritik/-verbesserung*	„Kritisch sein" wird als Selbstkritik mit dem Fokus der Selbstverbesserung verstanden.
	VK-NG: *Nicht direkt glauben*	„Kritisch sein" wird als Neigung definiert, nicht alles direkt zu glauben.
	VK-TK: *Tadeln/Korrigieren (anderer)*	„Kritisch sein" wird als das Tadeln anderer verstanden.
	VK-ND: *Nicht definierbar*	„Kritisch sein" kann nicht definiert werden oder es wird deutlich, dass keine klare Definition vorliegt.

Das *Verständnis der Selbstreflexion* (VS) wurde unter der Untersuchungsfrage 3

> Welches Verständnis haben Chemielehramtsstudierende vom Begriff der Selbstreflexion?

ausgewertet und in insgesamt sieben Kategorien unterteilt (Tabelle 44). Neben dem Verständnis von Selbstreflexion als das *Überdenken von Handlungen* (VS-HU) oder als *abwägender Prozess* (VS-AW), wurden hier ebenfalls Antworten, die keine klare Definitionsvorstellung erkennen ließen, der Kategorie *Nicht definierbar* (VS-ND) zugeordnet.

Tabelle 44: S2.I1 - Kategorien des *Verständnisses von Selbstreflexion* (VS)

Überkategorie	Kategorie	Definition
Verständnis Selbstreflexion (VS)	VS-HU: *Handlungen überdenken*	„Selbstreflexion" wird als das Überdenken (eigener) Handlungen verstanden.
	VS-IB: *Introspektive Betrachtung*	„Selbstreflexion" wird als introspektive Betrachtung verstanden.
	VS-SV: *Selbstverbesserung*	„Selbstreflexion" wird als Selbstverbesserung verstanden.
	VS-EM: *Empathie*	„Selbstreflexion" wird als Empathie verstanden.
	VS-AW: *Abwägen*	„Selbstreflexion" wird als abwägender Prozess verstanden.
	VS-HF: *Hinterfragen*	„Selbstreflexion" wird als das Hinterfragen von Dingen oder Themen verstanden.
	VS-ND: *Nicht definierbar*	„Selbstreflexion" kann nicht definiert werden oder es liegen diffuse Definitionsansätze vor.

Zur Einschätzung des *Fachwissens* (FW) (Tabelle 45) unter der Fragestellung 4

> Wie schätzen Studierende des Chemielehramtes ihr Fachwissen in dem gestellten Setting ein?

ergaben sich zwei Kategorien, wobei Aussagen das Fachwissen entweder als *unzureichend* (FW-UZ) oder als *grundlegend* (FW-GL) charakterisierten. *Grundlegend* (FW-GL) wurde kodiert, wenn die Proband*innen deutlich machten, dass ein Grundwissen zum zu bewertenden Kontext vorlag, welches aber nicht als umfangreich im Sinne von präzise oder konkret bezeichnet werden konnte. Die Kategorie *Unzureichend* (FW-UZ) wurde vergeben, wenn die Proband*innen angaben, kein Wissen zum erfragten Kontext zu haben.

Tabelle 45: S2.I1 - Kategorien zur *Einschätzung des Fachwissens* (FW)

Überkategorie	Kategorie	Definition
Einschätzung Fachwissen (FW)	FW-UZ: *Unzureichend*	Das eigene Fachwissen über den Kontext wird als unzureichend eingeschätzt.
	FW-GL: *Grundlegend*	Das eigene Fachwissen über den Kontext wird als grundlegend eingeschätzt.

Zur Beantwortung von Untersuchungsfrage 5

> Wie schätzen Studierende des Chemielehramtes ihre Bewertungsfähigkeit in Bezug auf den zu bewertenden Kontext ein?

wurden drei Kategorien gebildet (Tabelle 46). Die Befragten schätzten ihre Bewertungsfähigkeit entweder als *ausreichend* (BF-AU) ein, den Kontext zu bewerten, gaben an, den Kontext nur *eingeschränkt* (BF-EI) und unter Vorbehalt bewerten zu können oder sahen ihre Fähigkeit zur Bewertung als zu *unzureichend* (BF-UZ) an, um eine Bewertung vorzunehmen.

Tabelle 46: S2.I1 - Kategorien zur *Einschätzung der Bewertungsfähigkeit* (BF)

Überkategorie	Kategorie	Definition
Einschätzung Bewertungsfähigkeit (BF)	BF-AU: *Ausreichend*	Die eigene Fähigkeit zur Bewertung wird als ausreichend eingeschätzt.
	BF-EI: *Eingeschränkt*	Die eigene Fähigkeit zur Bewertung wird als eingeschränkt eingeschätzt.
	BF-UZ: *Unzureichend*	Die eigene Fähigkeit zur Bewertung wird als unzureichend eingeschätzt.

Für alle Untersuchungsfragen wurden die absoluten Häufigkeiten der Kategorien pro Proband*in in Excel aufgetragen und sowohl die Einfach- als auch die Mehrfachzählungen sowie der prozentuale Anteil einer Kategorie an den Gesamtkodierungen dargestellt.

5.2.1.2 Ergebnisse

Die Untersuchung der selbstreflexiven Aspekte ergab, dass die Teilnehmenden im Schnitt 3,2 verschiedene Aspekte während des Interviews reflektierten (Kategorie der *introspektiven Überprüfung*, SR-IU) (Tabelle 47). Hier wurde deutlich, dass die *Reflexion des eigenen Fachwissens* (FW) bei allen Proband*innen vorhanden war, was aber auch dem Umstand geschuldet war, dass im Interview explizit nach dieser Einschätzung gefragt wurde. Die Reflexion des eigenen Fachwissens äußerte sich beispielsweise in Äußerungen wie „So genau habe ich mich noch nie damit auseinandergesetzt“ (S2.I1-WI22-A-M03-CB33BT, Pos. 50) oder „kein fundiertes Vorwissen, also ich kann jetzt keine Quelle dazu nennen“ (S2.I1-WI22-B-B2_3-XF21FM, Pos. 61).

Tabelle 47: S2.I1 - Ergebnisse der Interviewerhebung, Selbstreflexion bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (SR), *N* = 10

Kategorie/Code		Summe Einfach-zählung	Summe Mehrfach-zählung	Summe rel. %	
SR-IU: Introspektive Überprüfung	SR-IU-FW: Bewertung des eigenen Fachwissens	10	23	**21,7**	**54,7**
	SR-IU-PR: Persönliche Relevanz (Vor- und Nachteile)	4	8	**7,5**	
	SR-IU-MW: Moral und Werte	4	6	**5,7**	
	SR-IU-RH: Retrospektive Handlungsbetrachtung	3	6	**5,7**	
	SR-IU-EE: Emotionale Einsicht	3	6	**5,7**	
	SR-IU-DG: Diffuses Gefühl	3	4	**3,8**	
	SR-IU-EF: Eigene Erfahrungen	1	1	**0,9**	
	SR-IU-ID: Identifikation	2	2	**1,9**	
	SR-IU-IN: Interesse	1	1	**0,9**	
	SR-IU-EN: Eigene Notwendigkeit	1	1	**0,9**	
Summe Introspektive Überprüfung (SR-IU)		**32**	**58**		
SR-EF: Bewusstsein für externe und interne Einfluss-faktoren	SR-EF-SN: Soziales Umfeld/Soziale Norm	9	23	**21,7**	**45,3**
	SR-EF-ME: Medien	6	8	**7,5**	
	SR-EF-PB: Persönliche Befangenheit/Vorurteile	4	4	**3,8**	
	SR-EF-IN: Interesse	3	4	**3,8**	
	SR-EF-ER: Kindheit und Erziehung	1	2	**1,9**	
	SR-EF-SY: Sympathie	2	2	**1,9**	
	SR-EF-BI: Bildung	1	1	**0,9**	
	SR-EF-PM: Persönlicher Möglichkeitsrahmen	1	1	**0,9**	
	SR-EF-EX: Expert*innenmeinungen	1	1	**0,9**	
	SR-EF-EF: Erfahrungen	1	1	**0,9**	
	SR-EF-MW: Moral und Werte	1	1	**0,9**	
Summe externe und intere Einflussfaktoren (SR-EF)		**21**	**25**		
Summe total		**53**	**83**		

Nicht explizierte Aspekte traten dagegen deutlich weniger in den Bewertungen der Befragten auf, was sich mit den Ergebnissen der Fragengenese S1.F1 (Kap. 5.1.1) und S.1.F2 (Kap. 5.1.2) deckte. Vier Proband*innen reflektierten die *persönliche Relevanz* (SR-IU-PR) in Aussagen wie „[…] was das halt für mich als Mensch bedeutet, ist es für mich persönlich teurer? Ist es für mich persönlich gut so?“ (S2.I1-WI22-B-

B2_3-BN10LB, Pos. 38) oder ihre *Moralvorstellungen und Werte* (SR-IU-MW), beispielsweise durch die Äußerung „[…] entspricht das man [sic] noch meinen Werten?“ (S2.I1-WI22-B-B2_3-BN10LB, Pos. 66). Die Kategorien der *retrospektiven Handlungsbetrachtung* (SR-IU-RH) und der *emotionalen Einsicht* (SR-IU-EE) machten jeweils 5,7 % aller Kodierungen im Bereich der *Selbstreflexion* (SR) aus und wurden von jeweils drei Teilnehmenden adressiert.

Ein *Bewusstsein für Einflussfaktoren* (SR-EF) zeigte sich vor allem im Bereich der *sozialen Norm* (SR-EF-SN) durch Äußerungen wie „Genauso würde ich aber auch im sozialen Umfeld sagen klar, wir lassen uns irgendwie immer unbewusst beeinflussen […]“ (S2.I1-WI22-B-B2_3-XF21FM, Pos. 86). Neun von zehn Befragten (90 %) machten so deutlich, dass ihr Umfeld eine relevante Wirkung auf sie selbst und ihre Meinungsbildung hat, was 21,7 % aller Kodierungen in dieser Überkategorie entsprach. Auch *Medien* (SR-EF-ME) wurden von sechs der zehn der Teilnehmenden (60 %) als beeinflussend wahrgenommen, wobei Werbung und die sozialen Medien zumeist hervorgehoben wurden. Beide Kategorien wurden im späteren Teil des Interviewleitfadens als expliziter Reflexionsstimulus erfragt, weshalb hier bereits von einer Häufung auszugehen war. Weitere Einflüsse, wie die *persönliche Befangenheit* (PB) (vier der zehn Proband*innen, 40 %) und das Interesse (IN) (drei der zehn Proband*innen, 30 %), wurden zwar seltener genannt, stellten jedoch mit die am häufigsten erwähnten Kategorien dar. Im Schnitt nahmen die Teilnehmenden 2,1 verschiedene Einflussfaktoren wahr. Die Verzerrung, die durch explizite Erfragung oder einen Reflexionsstimulus entstand, wurde in dieser Vorstudie erkannt und im weiteren Verlauf methodisch berücksichtigt. Sie trat sowohl bei der Einschätzung des Fachwissens als auch bei Reflexionsstimuli auf, die durch Beispiele angeregt wurden. In der darauffolgenden Hauptstudie wurde diese Verzerrung gezielt adressiert, indem ab Studie S3 sichergestellt wurde, dass Kodierungen eines Aspekts nicht mehr unmittelbar nach expliziten Nachfragen erfolgten. Derartige Nachfragen wurden zudem möglichst an das Ende des Interviews verlagert, um ihren Einfluss auf die Antworten zu minimieren.

Unter Rückbezug auf die Untersuchungsfrage 1 lässt sich konstatieren, dass ohne expliziten Stimulus wenig Selbstreflexion im Sinne einer introspektiven Überprüfung oder einer Berücksichtigung externer und interner Einflussfaktoren stattfand.

Mit Blick auf das Verständnis der Studierenden hinsichtlich der Begrifflichkeit *kritisch sein* (VK) zeigte sich ein vergleichbares Bild zu Studie S1. Sieben von zehn Befragten (70 %) verstanden darunter das *Hinterfragen von Dingen* (VK-HF) (Tabelle 48), wobei

die Aussagen dazu häufig unklar und unkonkret blieben, wie folgendes Beispiel zeigt: „Ja, dass man halt die Dinge nicht so nimmt, wie sie sind, sondern dass man wirklich hinterfragt, hinterfragt, was dahintersteckt." (S2.I1-WI22-B-B2_3-CB33FS, Pos. 86). Auch Umschreibungen wie „In dem Sinne, dass ich vielleicht die Situation manchmal auch nicht so akzeptiere, sondern immer noch Fragen stelle" (S2.I1-WI22-B-B2_3-XF21FM, Pos. 104) implizieren einen hinterfragenden Ansatz, ohne jedoch konkreter zu werden und etwa Kriterien zu nennen, die im Umgang mit Informationen erfüllt sein müssen.

Ein anderes Verständnis der Phrase *kritisch sein* zeigt sich in Zitaten wie „[…], dass ich halt erkenne, dass irgendwas nicht so mir gepasst hat, wie ich es gerne gehabt hätte, dass ich da eigentlich sehe, […] da hätte ich vielleicht irgendwas besser machen können […]" (S2.I1-WI22-A-M03-IV20FG, Pos. 66). Hier wird ein selbstbezogener Kritikbegriff deutlich, welcher eine prüfende Haltung sich selbst gegenüber beschreibt und sich als *Selbstkritik* (VK-SK) fassen lässt. 12,5 % der Kodierungen zu dieser Untersuchungsfrage fielen in diese Kategorie. Jeweils zwei Personen verstanden unter *kritisch sein* zudem eine *ungläubige Haltung* (VK-NG), das Kritik üben oder *Tadeln anderer* (VK-TK) oder konnten es für sich *nicht definieren* (VK-ND).

Tabelle 48: S2.I1 - Ergebnisse der Interviewerhebung, *Verständnis der Begrifflichkeit kritisch sein* (VK), *N* = 10

Kategorie/Code	Summe Einfach-zählung	Summe Mehrfach-zählung	Summe rel. %
VK-HF: Hinterfragen	7	12	**50,0**
VK-WW: Widerspricht eigenen Werten	1	3	**12,5**
VK-SK: Selbstkritik/ -verbesserung	3	3	**12,5**
VK-NG: Nicht direkt glauben	2	2	**8,3**
VK-TK: Tadeln/Korrigieren	2	2	**8,3**
VK-ND: Nicht definierbar	2	2	**8,3**

Das *Verständnis von Selbstreflexion* (VS) lag breit gefächert vor. Drei der Befragten gaben an, dass der Begriff für sie *nicht definierbar* (VS-ND) sei (Tabelle 49), wohingegen weitere drei darunter das *Überdenken der eigenen Handlungen* (VS-HU) verstanden. Stellvertretend dafür kann die Äußerung „reflektiert ist, glaube ich, vor allen Dingen, dass ich auch im Nachhinein Sachen überdenke" (S2.I1-WI22-A2-M03-MP17JN, Pos. 66) angeführt werden. Diese Betrachtung eigener Handlungen war bei zwei Proband*innen zudem mit der *eigenen Verbesserung* (VS-SV) verknüpft, was sich in der

Äußerung „Aber nicht nur noch mal generell drüber nachdenken, sondern auch immer zu gucken, wie kann ich nochmal vielleicht was verbessern oder was habe ich auch gut gemacht oder schlecht natürlich auch“ (S2.I1-WI22-B-B2_3-XF21FM, Pos. 98) widerspiegelt. Die Selbstreflexion wurde hingegen nur von drei der zehn Teilnehmenden (30 %) als *introspektive Betrachtung* (VS-IB) verstanden, was sich in exemplarisch formulierten selbstreflexiven Fragen, wie beispielsweise „[…] wie empfinde ich gerade?“ oder „[…] wie habe ich mich da gefühlt“ (S2.I1-WI22-B-B2_3-BN10LB, Pos. 66), niederschlug.

Tabelle 49: S2.I1 - Ergebnisse der Interviewerhebung, *Verständnis des Begriffs Selbstreflexion* (VS), *N* = 10

Kategorie/Code	Summe Einfach-zählung	Summe Mehrfach-zählung	Summe rel. %
VS-HU: Handlungen überdenken	3	4	**21,1**
VS-IB: Introspektive Betrachtung	3	5	**26,3**
VS-SV: Selbstverbesserung	2	4	**21,1**
VS-EM: Empathie	1	1	**5,3**
VS-AW: Abwägen	1	1	**5,3**
VS-HF: Hinterfragen	1	1	**5,3**
VF-ND: Nicht definierbar	3	3	**15,8**

Die *Einschätzung des eigenen Fachwissens* (FW) erfolgte durch die Teilnehmenden sowohl selbstständig als auch auf explizite Nachfrage. Die Mehrheit (acht von zehn Teilnehmenden, 80 %) gab dabei mindestens einmal im Interview an, nur u*nzureichendes Wissen* (FW-UZ) zum Kontext zu besitzen (Tabelle 50). Auffällig war, dass fünf Teilnehmende im Laufe des Interviews sowohl angaben, nur über *unzureichendes Fachwissen* (FW-UZ) zu verfügen, als auch deutlich machten, ein *grundlegendes Verständnis* (FW-GL) des Kontextes zu haben, was auf eine Unsicherheit bezüglich ihres Wissensfundus zurückzuführen sein könnte.

Beispielsweise bezeichnete sich eine Person im Kontext des Interviews als „Laie“ (S2.I1-WI22-A2-B2_3-XJ27BO, Pos. 28) und erklärte auf spätere Nachfrage, dass sie „die Basics“ (Pos. 38) kenne, jedoch nicht über genügend Wissen verfüge, um sich als Expert*in zu identifizieren. Während die erste Aussage eine Einschätzung des Fachwissens als *unzureichend* (FW-UZ) nahelegte, deutete die zweite eher auf eine Bewertung als *grundlegend* (FW-GL) hin. Ein vergleichbares Beispiel einer anderen interviewten Person ist die Aussage, dass sie ihr Fachwissen auf einer „[…] Skala von

1 bis 5 […]“ (S2.I1-WI22-A1-M03-CB33BT, Pos. 40) bei zwei einordnen würde, im späteren Verlauf aber angab, „[…] kein fundiertes Vorwissen […]“ (Pos. 61) zu haben.

Daraus ergab sich, dass ebenfalls sieben Proband*innen mindestens einmal in Interview ihr Fachwissen als *grundlegend* (FW-GL) in Bezug auf den zu bewertenden Kontext einschätzten. Eine klare Selbsteinschätzung, die auf ein umfangreiches Grundwissen schließen ließ, fand bei keiner Person statt.

Tabelle 50: S2.I1 - Ergebnisse der Interviewerhebung, *Selbsteinschätzung des Fachwissens* (FW), *N* = 10

Kategorie/Code	Summe Einfach-zählung	Summe Mehrfach-zählung	Summe rel. %
FW-ZU: Unzureichend	8	11	**57,9**
FW-GL: Grundlegend	7	8	**42,1**

Neben der *Einschätzung des eigenen Fachwissens* (FW) war im Interviewkontext auch die *Einschätzung der Bewertungsfähigkeit* (BF) von Interesse (Tabelle 51). Zwei Teilnehmende sahen sich *ausreichend* (BF-AU) befähigt, um eine Bewertung vornehmen zu können, während fünf Teilnehmende sich zumindest *eingeschränkt* (BF-EI) in der Lage sahen, den Kontext zu bewerten, wie folgendes Zitat exemplarisch zeigt: „Ich glaube, eine grundlegende Meinung ja, aber noch nicht tiefgreifend genug, um eine feste Entscheidung jetzt zu fällen“ (S2.I1-WI22-A2-M03-MP17JN, Pos. 28). Nur drei sahen sich als *unzureichend* (BF-UZ) informiert an und lehnten auf dieser Grundlage eine Positionierung ab. Mit Rückblick auf die Ergebnisse von Untersuchungsfrage 4 fällt auf, dass, obwohl bezüglich des Wissens zum Kontext eine große Unsicherheit offensichtlich wurde, sich sieben Befragte zumindest unter Einschränkungen in der Lage sahen, den Kontext zu bewerten und sich positionieren zu können.

Tabelle 51: S2.I1 - Ergebnisse der Interviewerhebung, *Selbsteinschätzung der Bewertungsfähigkeit* (BF), *N* = 10

Kategorie/Code	Summe Einfach-zählung	Summe Mehrfach-zählung	Summe rel. %
BF-AU: Ausreichend	2	2	**20,0**
BF-EI: Eingeschränkt	5	5	**50,0**
BF-ZU: Unzureichend	3	3	**30,0**

Eine besondere Aussagekraft erhält eine kombinierte Betrachtung beider Kategorien zur *Selbsteinschätzung Fachwissen* (FW) und der *Bewertungsfähigkeit* (BF), insbesondere

dann, wenn beide Einschätzungen Inkongruenzen aufweisen. Als Beispiel sei ein Zitat aus der Erhebung S2.I1 angeführt:

> I: Wieviel weißt du denn bereits über das Thema? Du hast ja schon einiges genannt, aber würdest du sagen, dass es für dich ausreichend ist? #00:08:21-0#
>
> B: Ja, man könnte schon mehr wissen, aber es ist halt so das, was man so was man so erfährt. Also (unv.) mein Gott, klar, das sind Siliziumzellen, das weiß man. Dann gibt es das Silicon Valley (lacht). Na gut, das sind so banale Dinge. Aber, ja so viel weiß ich natürlich nicht. Ich weiß nicht genau, wie die die Technik ist das alles so ein Brimborium, was mich nicht so ganz so interessiert. #00:08:42-1#
>
> I: Aber du würdest schon sagen, dass diese diese [sic] Menge an Wissen ausreicht für dich, um diese Entscheidung treffen zu können. #00:08:47-3#
>
> B: Ja sicherlich. #00:08:48-2# #00:08:49-6#
> (S2.I1-Wi22-A-B2_3-CB33FS, Pos. 49-52)

Dieses Zitat verdeutlicht ein Bewusstsein für den Mangel an Wissen, während gleichzeitig ein subjektives Gefühl der Fähigkeit zur Bewertung besteht. Diese Verzerrung zeigte sich bei zwei Teilnehmenden besonders deutlich, was sich darin äußerte, dass sie ihr eigenes Fachwissen als *unzureichend* (FW-UZ), ihre Bewertungsfähigkeit hingegen als *ausreichend* (BF-AU) einschätzten. Sechs weitere Befragte gaben mindestens einmal im Interview an, *unzureichendes Fachwissen* (FW-UZ) zum Kontext zu haben, sahen sich aber *eingeschränkt* (BF-EI) in der Lage, sich zu positionieren. Dennoch bleibt die Frage ungeklärt, ob die Proband*innen grundsätzlich in der Lage sind, naturwissenschaftliche Fragen zur Aneignung von Wissen zu formulieren, was einen Einfluss auf das subjektive Empfinden der Bewertungsfähigkeit haben könnte.

5.2.2 Fragebogenerhebung S2.F1

Der in der Studie S2 eingesetzte Fragebogen bestand aus zwei bereits etablierten und auf Reliabilität geprüften Fragebögen, die im Zuge der Erhebung miteinander kombiniert wurden. Einmal dem CRT nach Frederick (2005) und dem SRIS nach Silvia (2021), welcher auf dem ursprünglichen SRIS von Grant et al. (2002) basiert. Eine detailliertere Betrachtung kann in Kapitel 4.2.3 nachgelesen werden.

Die Items beider Fragebögen wurden im hier verwendeten Fragebogen so angeordnet, dass die drei CRT-Fragen und die beiden SRIS-Blöcke abwechselnd den Teilnehmenden gestellt wurden, wobei die Reihenfolge der CRT-Fragen und SRIS-Blöcke durch LimeSurvey randomisiert erfolgte. Alle Teilnehmenden der Interviewerhebung nahmen auch an der hier dargestellten Fragebogenerhebung teil, weshalb die Stichprobe $N = 10$ betrug.

5.2.2.1 Auswertung

Zur Auswertung wurden die schriftlich formulierten Stufen der Likert-Skala der SRIS in numerische Werte umgewandelt und die Items IN 2, 4, 5, 6 und 7, die negativ formuliert waren, invers kodiert. Die Ergebnisse der Items wurden tabellarisch dargestellt und, unter der Annahme einer quasi-metrischen Skalierung (Urban & Mayerl, 2011; Völkl & Korb, 2018) (vgl. Kap. 4.3.2), mittels deskriptiver Statistik ausgewertet. Besonders relevant bei der Auswertung der SRIS sind die Mittelwerte der Subskalen *Selbstreflexion* (SR) und *Einsicht* (IN). Diese ermöglichen insbesondere im Vergleich mehrerer Proband*innen eine differenzierte Analyse individueller Ausprägungen sowie gruppenübergreifender Muster.

Der CRT wurde analog zu Studie S1.F3 ausgewertet (vgl. Kap. 5.1.3.1).

5.2.2.2 Ergebnisse

Die Auswertung der SRIS zeigte eine hohe Streuung der Daten innerhalb der Gruppe, was auf deutliche individuelle Unterschiede in den Neigungen zur Selbstreflexion und Einsicht schließen lässt (Tabelle 52). Dies deutet darauf hin, dass die Teilnehmenden unterschiedliche Herangehensweisen an introspektive Prozesse aufweisen. Bei einigen Proband*innen zeigte sich dadurch eine höhere Reflexionsbereitschaft und ein besseres Verständnis der eigenen Denkprozesse als bei anderen (vgl. Tabelle 53).

Da der mittlere Wert einer siebenstufigen Likert-Skala bei 4 liegt und als neutral in Bezug auf die Zustimmung definiert wird, sind die Mittelwerte von 5,3 bei der *Selbstreflexion* (SR) und 4,7 bei der *Einsicht* (IN) als leicht positive Tendenzen festzuhalten, selbstreflexives Verhalten zu zeigen und daraus auch Erkenntnisse gewinnen zu können (Tabelle 53). Grundsätzlich fällt der positive Trend im Bereich der *Selbstreflexion* (SR) stärker aus als bei der *Einsicht* (IN), was nahelegt, dass die Bereitschaft zur Selbstreflexion zwar vorhanden ist, die introspektiv gewonnenen Erkenntnisse aber nicht immer eingeordnet oder verstanden werden können.

Tabelle 52: S2.F1 - Ergebnisse der SRIS nach Silvia (2021), *N* = 10

Code	SR3	SR5	SR6	SR8	SR9	SR10	Summe SR	Summe SR rel. %
Summe	53	52	52	53	56	52	318	75,7
MW	**5,300**	**5,200**	**5,200**	**5,300**	**5,600**	**5,200**	31,800	75,7
Stabw.	1,567	1,476	1,874	1,703	1,506	1,398	6,925	16,5
Varianz	2,456	2,178	3,511	2,900	2,267	1,956	47,956	271,9
Modus	7	5	7	7	7	6	36	85,7
Median	5,5	5	5,5	6	6	6	34	81,0
Min	3	2	2	3	3	3	18	42,9
Max	7	7	7	7	7	7	40	95,2

Code	IN2	IN4	IN5	IN6	IN7	IN8	Summe IN	Summe IN rel. %
Summe	42	44	52	40	51	50	279	66,4
MW	**4,200**	**4,400**	**5,200**	**4,000**	**5,100**	**5,000**	27,900	66,4
Stabw.	1,687	1,578	1,751	1,886	1,595	1,414	6,173	14,7
Varianz	2,844	2,489	3,067	3,556	2,544	2,000	38,100	216,0
Modus	6	5	6	4	6	5	28,000	66,7
Median	4,5	5	6	4	6	5	28,000	66,7
Min	1	2	2	1	2	2	19,000	45,2
Max	6	6	7	7	7	7	36,000	85,7

MW SR	MW IN	MW Gesamt	Stabw. SR	Stabw. IN	Stabw. Ges.
Gesamt SR	Gesamt IN	Gesamt SR+IN	Gesamt SR	Gesamt IN	Gesamt SR+IN
5,300	**4,650**	**4,975**	1,533	1,655	1,622

Tabelle 53: S2.F1 - Ergebnisse der SRIS nach Silvia (2021), Mittelwerte pro Proband*in, *N* = 10

Code	MW SR	MW IN	MW Gesamt
MP17JN	6,000	5,667	5,833
XJ27BO	4,833	6,000	5,417
DB33BO	6,000	5,833	5,917
CB33BT	5,500	3,667	4,583
CB33FS	3,000	3,167	3,083
IV20FG	4,333	4,333	4,333
XJ36JE	5,833	4,667	5,250
NÜ20FG	4,333	3,333	3,833
XF21FM	6,500	4,667	5,583
BN10LB	6,667	5,167	5,917
	Gesamt SR	**Gesamt IN**	**Gesamt SR+IN**
	5,300	**4,650**	**4,975**

Vier Teilnehmende wiesen zudem deutlich hohe Werte (> 6), insbesondere in der Subskala der *Selbstreflexion* (SR), auf. Diese Werte liegen dabei nicht nur über dem neutralen Skalenmittelpunkt von 4, sondern auch deutlich über dem Gruppendurchschnitt von 5,3. Dies deutet auf eine ausgeprägte Selbstreflexionsneigung hin, wobei zu berücksichtigen ist, dass Selbsteinschätzungen potenziell durch Verzerrungen zugunsten eines positiven Selbstbildes beeinflusst sein können (Dunning et al., 2002). Dennoch ist in beiden Blöcken ein leicht positiver Trend erkennbar. Insgesamt lässt sich daher festhalten, dass eine Bereitschaft zur Selbstreflexion vorhanden zu sein scheint, auch wenn diese nicht als stark ausgeprägt beschrieben werden kann.

Die Ergebnisse des CRT wiesen eine Normalverteilung der Ergebnisse hinsichtlich der Anzahl korrekter Antworten und damit eine Vergleichbarkeit mit den Referenzdaten von Frederick (2005) und den Ergebnissen der Studie S1.F3 (vgl. Kap. 5.1.3.2) auf. Im Schnitt beantworteten die Teilnehmenden 1,5 Fragen korrekt (Tabelle 54). Die inkorrekten Antworten entsprachen auch hier dem erwarteten Muster der intuitiven Erstantworten. Vergleichbar zu Studie S1.F3 liegt auch hier der Schluss nahe, dass bei 50 % der Gruppe eine eingeschränkte Erstimpulsinhibition zugunsten heuristischer Urteile vorliegt.

Tabelle 54: S2.F1 - Ergebnisse des CRT nach Frederick (2005), *N* = 10

Code	CRT1	CRT2	CRT3	Korrekt	%
Korrekt	3	7	5	15	**50,0**

Verteilung der Ergebisse in %			
20,0	30,0	30,0	20,0
0	1	2	3
Low			High

5.2.3 Abschluss der Studie S2

Die Erhebungsstruktur der Studie S2 ermöglichte erstmals eine umfassendere Kombination aus der Selbst- und Fremdeinschätzung hinsichtlich der Selbstreflexion.

Zusammenfassend zeigte sich im Rahmen der Fremdeinschätzung, dass in den verwendeten Bewertungskontexten der Interviewerhebung nicht explizierte Aspekte der Selbstreflexion deutlich untergeordnet auftraten und auch Einflussfaktoren selten ohne Stimulus reflektiert wurden. In der Lebensrealität sind explizite Stimuli jedoch selten, was die Notwendigkeit einer selbstständigen Reflexionsfähigkeit unterstreicht.

Zudem lag die Phrase *kritisch sein* wenig präzise definiert bei den Teilnehmenden vor. Das mehrheitliche Verständnis von *kritisch sein* als hinterfragen oder nicht glauben

deutet in diesem Kontext auf eine grundsätzlich skeptische Haltung hin, wobei Skepsis allein jedoch nur wenig praktische Anweisungen für alltägliche Entscheidungen bietet, vor allem, wenn diese nicht präzisiert und zielgerichtet vorliegt (Collins, 2009; Sharon & Baram-Tsabari, 2020). Ähnlich verhielt es sich mit dem Verständnis von Selbstreflexion, welches ebenfalls subjektiv geladen und in Folge breit gefächert vorlag.

Das Fachwissen innerhalb der Kontexte wurde größtenteils als unzureichend und maximal als grundlegend eingeschätzt, wohingegen sich der Großteil der Befragten zumindest unter Vorbehalt in der Lage sah, eine Bewertung vornehmen zu können. Wenn das Wissen aber als Grundlage für eine Positionierung als nicht ausreichend angesehen werden kann, müssen dem folgend andere Aspekte als relevant erachtet worden sein, welche das subjektive Gefühl der Entscheidungsfähigkeit vermittelten. Hier liegt die Annahme nahe, dass das intuitive System 1 auf Basis unterbewusster Hinweisreize (Intuition) ein Gefühl für eine Positionierung erzeugte, welches auch trotz des geringen Wissens fortbestand. Dies würde sich mit den Ergebnissen aus S1.F1 und S1.F2 zur Stärke des intuitiven Systems in Entscheidungsprozessen decken und verdichtete die Annahme, dass eine Adressierung der affektiv intuitiven Ebene als relevant angesehen werden kann. Dennoch konnte auf Grundlage der bisherigen Fragestellungen nicht ausgeschlossen werden, dass nicht auch die Fähigkeit, naturwissenschaftliche Fragestellungen zu formulieren, welche dann zur selbstständigen Wissensaneignung führen, zu einer empfundenen Bewertungsfähigkeit beiträgt, obwohl noch kein fachliches Wissen zum Kontext vorhanden ist. Infolgedessen erschienen die Anpassung und Erweiterung der Untersuchungsfragen hinsichtlich der Aspekte naturwissenschaftlicher Bewertung als sinnvolle ergänzende Schritte zur weiteren Ausdifferenzierung.

Dennoch zeigt auch der CRT bei der Hälfte der Teilnehmenden Hinweise auf eine schwache Erstimpulsinhibition und infolgedessen eine verstärkte Bewertung auf Basis heuristischer Urteile.

Im Gegensatz dazu bewerteten die Befragten im SRIS ihre Tendenz, selbstreflexives Verhalten zu zeigen, überwiegend positiv, während die daraus gewonnenen Einsichten etwas schwächer eingeschätzt wurden. Daraus wächst ein Spannungsgefüge zwischen den deutlich geringen selbstreflexiven Anteilen bei der Bewertung der Interviewkontexte und der Selbsteinschätzung bezüglich selbstreflexiven Verhaltens, dem in weiteren Studien dieses Projekts nachgegangen werden soll.

Aus methodischer Sicht zeigten sich im Interview keine erkennbaren qualitativen Unterschiede zwischen den Ergebnissen der beiden Kontexte A1 (Solarenergieversorgung vs. Windkraftanlagen) und B (E-Mobilität vs. Brennstoffzellenantrieb). Beide zeigten sich für die Zielsetzung als geeignet, weshalb sie in weiteren Erhebungen Anwendung fanden. Der Kontext A2, der die Energiegewinnungsformen Atomenergie und Kohleverbrennung umfasste, erwies sich hingegen als wenig ergiebig, da eine direkte Ablehnung der kontrovers diskutierten Energieformen erfolgte, was zulasten einer tiefergehenden inhaltlichen Auseinandersetzung ging. In Konsequenz wurde der Kontext A2 in weiteren Erhebungen nicht verwendet.

Die Kodierungen von Antworten auf explizite Nachfragen oder Trigger, wie im Falle des *Fachwissens* (SR-IU-FW) bei den Aspekten der *Selbstreflexion* (SR), wurden in nachfolgenden Studien angepasst. Derlei Fragen wurden ans Ende des Interviews verlagert, um mögliche Interferenzen zu minimieren, und es wurden infolgedessen keine Kodierungen von Antworten vorgenommen, die direkt auf explizite Nachfragen zu Aspekten der Selbstreflexion zurückzuführen waren.

Die SRIS und der CRT erwiesen sich als sehr praktikable ergänzende Instrumente, weshalb sie in den nächsten Studien ebenfalls eingesetzt wurden.

Die Erkenntnisse aus Studie S1 und Studie S2 zeigten einen starken Einfluss der intuitiv-affektiven Ebene auf, welche in Entscheidungsprozessen die kognitiv-rationale Ebene beeinflusst oder ersetzt. Dies legt den Schluss nahe, dass die Annahmen der Dual-Process Theory (vgl. Kap. 2.3.2; Haidt, 2008; Kahneman, 2012; Osman, 2004; Thompson et al., 2011, 2013) auch bei den beobachteten Stichproben zugrunde gelegt werden können. Ansatzpunkte hierfür ließen sich sowohl in den durchgeführten Fragebogenformaten aus S1.F1-3 als auch in den Interviews aus S2.I1 erkennen.

Hinzu kommt, dass die Begrifflichkeiten der Selbstreflexion und des kritisch Seins wenig sinnerfüllt und, wenn überhaupt, subjektiv geladen vorliegen, was einer zielgerichteten und effektiven Vermittlung dieser Konzepte durch Lehramtskandidat*innen in ihrer Rolle als Bildungsmultiplikator*innen entgegensteht. Die grundsätzliche Fähigkeit, naturwissenschaftliche Fragen zu stellen, scheint vorhanden zu sein, ohne eine explizite Aufforderung von außen kommt diese jedoch kaum zum Tragen. In Anbetracht des Umstands, dass in der Lebensrealität selten explizite Stimuli vorhanden sind, wird die Notwendigkeit einer zielgerichteten Sensibilisierung und Förderung deutlich.

Eine *explizite* Thematisierung erscheint dabei als weiterer notwendiger Schritt, was sich mit den Erkenntnissen aus der Vermittlung ähnlich komplexer Konstrukte, wie NOS, deckt (Khishfe & Abd-El-Khalick, 2002; Lederman & Lederman, 2014). In Studie S3 sollte folglich das Instrument zur Operationalisierung Kritischen Denkens durch den Einsatz einer Vermittlungseinheit, welche anstrebt, Kritisches Denken im Sinne eines multiperspektivischen Denkansatzes sowie Selbstreflexion als dessen konstitutives Element zu fördern, auf Funktionalität geprüft werden.

5.3 Studie S3

Die im Sommersemester 2023 durchgeführte Studie S3 bildete auf Grundlage der vorangegangenen Vorstudien die erste Hauptstudie des Projekts. In ihr flossen die methodischen Erkenntnisse der Vorerhebungen mit den währenddessen entwickelten theoretischen und konzeptionellen Ansätzen zusammen, die in Kapitel 3 bereits vorgestellt wurden.

Die auf diesen konzeptionellen Ansätzen basierende Definition Kritischen Denkens (vgl. Kap. 3.3) bildete eine relevante Grundlage für die spätere Kategorisierung und wird deshalb hier erneut angeführt:

> Kritisches Denken lässt sich charakterisieren als eine sorgfältige, systematische und vor allem begründete Entscheidungsfindung bei der Auseinandersetzung der*s Denkenden mit einem Objekt oder Subjekt. Diese Auseinandersetzung beinhaltet die Evaluation und Anwendung des eigenen Wissens und der eigenen Fähigkeiten auf kognitiver Ebene sowie die Berücksichtigung der eigenen Emotionen, Werte, Einstellungen und Motivationen auf affektiver Ebene. Sie umfasst zudem eine konstante Selbstreflexion und -regulation, die eine Rückkopplung zur*m Denkenden selbst ermöglicht und diese*n nicht nur in die Position der*s Bewertenden, sondern auch in die Position der*s zu Bewertenden versetzt (eigene Definition s. Kap. 3.3 u. a. in Anlehnung an Haidt, 2008; Halpern, 2014; Kahneman, 2012; D. Kuhn, 1999; Rafolt et al., 2019).

Ergänzend dazu und zur Annäherung an die affektive Ebene und die damit verbundenen Aspekte, wie Werte, Normen, Haltungen, Motivationen und Emotionen, wurde auch der Begriff der Selbstreflexion definierend gefasst:

> Die Selbstreflexion ist ein Prozess der introspektiven Überprüfung eigener Gedanken, Gefühle, Verhaltensweisen und Prägungen (Motivation, soziales Umfeld, eigene Erfahrungen u. v. m.) sowie darauf aufbauend ein Bewusstsein für

interne und externe Einflussfaktoren. Es umfasst die Wahrnehmung, Gewichtung und Bewertung ebenjener, um diese in Entscheidungsprozessen berücksichtigen zu können und so eine objektivere Position zu erhalten (eigene Definition s. Kap. 3.3 in Anlehnung an Carver & Scheier, 2001; Facione, 1990; Gigerenzer, 1991; Grant et al., 2002; Silvia, 2021).

Diese Definitionen und die damit verbundenen konzeptionellen Überlegungen dienten im weiteren Verlauf als zentrale Leitlinien bei der Erhebung und Auswertung der Daten, wobei das Triadische Modell Kritischen Denkens in der Naturwissenschaftsdidaktik (s. Abbildung 13 in Kap. 3.5) eine Schlüsselrolle einnahm. Die Teilbereiche des Modells, das naturwissenschaftliche Wissen, das epistemologische Verständnis und nicht zuletzt die Selbstreflexion, spiegelten im weiteren Verlauf den übergeordneten Rahmen wider, in dem die Kategorisierung und damit Operationalisierung des Kritischen Denkens, insbesondere in den Interviewerhebungen, erfolgte.

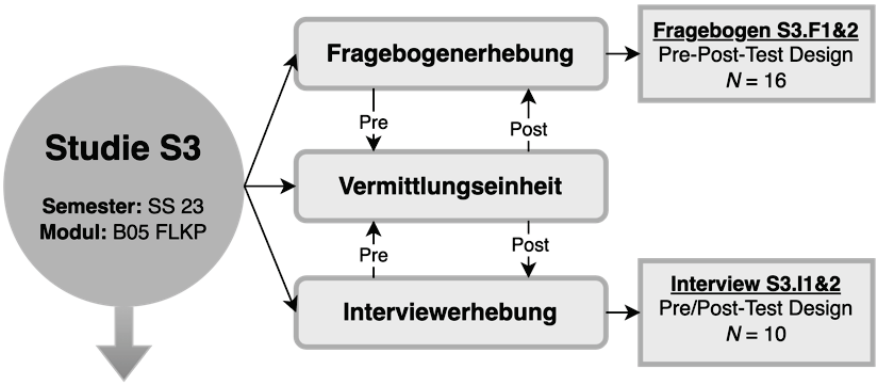


Abbildung 20: Übersicht über den Aufbau und die Erhebungsparameter der Studie S3

Die Studie S3 bestand aus einer Interview- und einer Fragebogenerhebung, um weiterhin dem Anspruch und dualen Ansatz der Verbindung aus Fremd- und Selbsteinschätzung Rechnung zu tragen. Durchgeführt wurden beide Erhebungen im *Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP)*, welches vom Untersuchenden als Lehrperson gestaltet wurde (vgl. Kap. 4.2.2). Neben Studierenden, die bisher noch an keiner der Befragungen teilgenommen hatten, nahmen auch vier Teilnehmende aus der vorangegangenen Studie S2 am Seminar und den damit verbundenen Erhebungen teil (Abbildung 20).

Das in Studie S3 erstmals zum Einsatz gebrachte neue Fragebogenformat wurde im Vorfeld mit $N = 4$ Studierenden aus dem chemiedidaktischen Mastermodul *Vertiefende Aspekte der Chemiedidaktik (M1 HR/SP/GG)* pilotiert und zusätzlich mit einem Doktoranden der Chemiedidaktik durchgeführt (vgl. Kap. 4.2.2). Dies diente zum

einen dem überprüfenden Ersteinsatz mit Teilnehmenden, die in ihren Voraussetzungen der späteren Stichprobe ähnelten, als auch der Überprüfung durch eine Person mit empirischer Expertise.

Insgesamt nahmen $N = 16$ Personen an der Fragebogenstudie teil, von denen $N = 10$ zusätzlich an der Interviewstudie partizipierten.

5.3.1 Interviewerhebung S3.I1&2

Die Interviewerhebung S3.I1&2 stellte zur Konsolidierung des vorangegangenen Interviewformats eine Wiederholung und Ausschärfung dieses Testinstruments dar. Dies äußerte sich im erneuten Einsatz der beiden bereits erfolgreich angewendeten Interviewleitfäden A und B (Anhang A.1.1) sowie den damit verbundenen Ausgangslagen für die Interviews (s. Kap. 4.2.6.2 & 4.2.6.3). Der Kontext A2, der zusätzlich zu den beiden benannten erneuerbaren Energiequellen auch die Atomenergie und Kohleverbrennung einbezog, wurde auf Grundlage der in Studie S2 gewonnenen Erkenntnisse nicht weiterverwendet (vgl. Kap. 5.2.3).

Im Gegensatz zur Studie S2 wurde zudem die initiale Pre-Test-Erhebung nach Abschluss des Seminars um eine Post-Test-Erhebung komplettiert, um das Instrument auf Funktionalität zu testen und etwaige Veränderungen in den Bewertungsaspekten der Studierenden aufzuzeigen. Die dazwischenliegende Seminareinheit inklusive der Methode der Systemisch-reflexiven Stellungnahme kann detailliert in Kapitel 4.2.5 nachgelesen werden. Die Post-Tests behandelten stets den jeweils anderen und nicht im Pre-Test verwendeten Kontext, wobei der Startkontext den Studierenden derart randomisiert zugeteilt wurde, dass gleich viele Teilnehmende mit Kontext A wie mit Kontext B im Pre-Test begannen.

Da vier der Studierenden des Seminars bereits bei der Interviewerhebung S2.I1 der Studie S2 teilgenommen hatten, welche als Pre-Test deckungsgleich in Leitfaden und Setting mit dem hier in Studie S3.I1&2 durchgeführten Interviewformat war, ergab sich die Möglichkeit, diese Studierenden zu einem Post-Interview zu bitten, an dem alle vier Personen freiwillig teilnahmen. Die Gesamtstichprobe der Interviewerhebung S3.I1&2 umfasste daher $N = 10$ Teilnehmende, bestehend aus sechs neu teilnehmenden Studierenden und vier weiteren Teilnehmenden, mit denen ein ergänzendes Post-Interview abgehalten wurde.

Die Erhebung der Pre-Interviews erfolgte in der ersten und zweiten Woche des Seminars, bevor thematisch relevante Inhalte innerhalb des Seminars thematisiert wurden (vgl. Kap. 4.2.5). Die Post-Interviews fanden je nach Verfügbarkeit der Teilnehmenden etwa zwei bis vier Wochen nach Abschluss des Seminars statt.

Zusätzlich zu den Post-Interviews wurden mit drei Freiwilligen nach etwa sechs Monaten Follow-Up-Interviews geführt, um die Langfristigkeit der Ergebnisse einschätzen zu können.

5.3.1.1 Auswertung

Die Auswertung der Interviews erfolgte rekurrierend auf die konzeptionellen Ansätze bezüglich der Definition und Bedeutung Kritischen Denkens in der Naturwissenschaftsdidaktik (vgl. Kap. 3). Das Triadische Modell Kritischen Denkens, welches ob seiner Relevanz hier nochmals abgebildet wird (s. Abbildung 21 oder Abbildung 13 in Kap. 3.5), stellte dabei die drei großen Teilbereiche, die als Grundlage für die weitere Systematisierung und Strukturierung der Untersuchungen sowie als deduktive Leitlinie dienten.

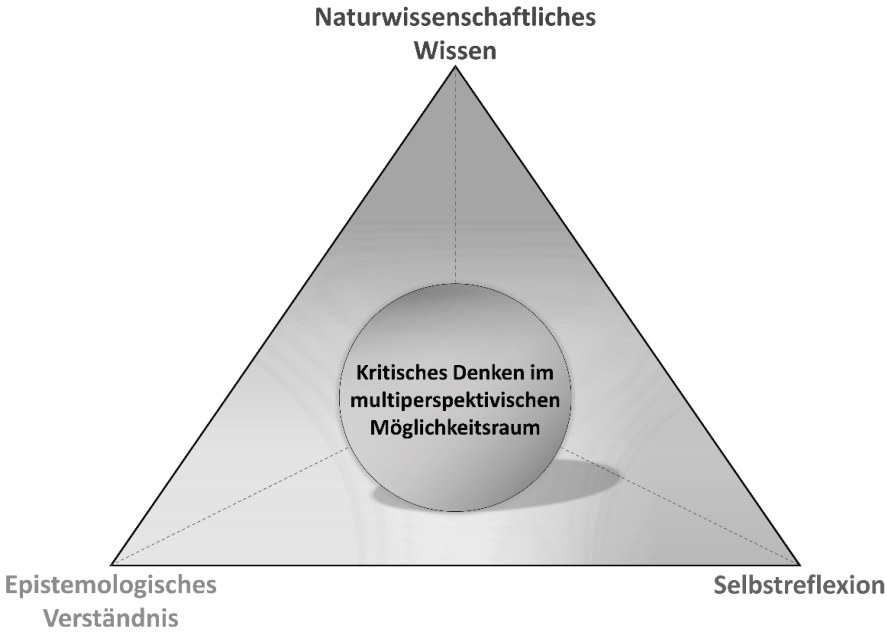


Abbildung 21: Triadisches Modell Kritischen Denkens in der Naturwissenschaftsdidaktik, näher erläutert in Kapitel 3.5

Motiviert durch die Ergebnisse der Studien S1 und S2 wurde eine deutlich stärkere Differenzierung und Systematisierung der Untersuchungsfragen vorgenommen und ergänzend zur *Selbstreflexion* (SR-A) auch das *epistemologische Verständnis* (EP-A) und das *naturwissenschaftliche Wissen* (NW-A) in den Bewertungsprozessen untersucht (Tabelle 55). Dies diente in Anlehnung an die noch im Fragebogenformat durchgeführte

Fragengenese in Studie S1 (vgl. Kap. 5.1) der Ermittlung der Denkperspektiven in konkreten Kontexten mit naturwissenschaftlichem Bezug, um eine vom tatsächlichen Fachwissen unabhängige Betrachtung zu ermöglichen und somit auch das Potenzial, sich Wissen anzueignen, zu adressieren. Zusätzlich dazu wurden die Untersuchungsfragen zum Verständnis der Begrifflichkeiten *Selbstreflexion* (SR-V) und *Kritisches Denken* (KD-V) beibehalten. Abschließend wurden die *Selbsteinschätzung des Fachwissens* (NW-SF) und die tatsächliche *Anwendung von Fachwissen* (NW-FB) im Bewertungskontext sowie die empfundene *Bewertungsfähigkeit der Teilnehmenden* (FB) als ergänzende Informationen beleuchtet, auch wenn diese nicht im Hauptfokus der Untersuchung standen.

Tabelle 55: S3.I1&2 - Konzeptionelle Teilbereiche und zugehörige Untersuchungsfragen

Teilbereich	Zugehörige Untersuchungsfrage
Naturwissenschaftliches Wissen (NW)	– Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug fachliche Aspekte? (NW-A) – Inwieweit findet Fachwissen bei Chemielehramtsstudierenden Einzug in die Bewertung des Kontextes? (NW-FB) – Wie schätzen Chemielehramtsstudierende ihr Fachwissen in dem gestellten Kontext ein? (NW-SF)
Selbstreflexion (SR)	– Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug selbstreflexive Aspekte? (SR-A) – Welches Verständnis haben Chemielehramtsstudierende von dem Begriff der Selbstreflexion? (SR-V)
Epistemologisches Verständnis (EP)	– Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug epistemologische Aspekte? (EP-A)
Kritisches Denken (KD)	– Welches Verständnis haben Chemielehramtsstudierende von der Formulierung *kritisch denken* oder *kritisch sein*? (KD-V)
Fähigkeit zur Bewertung (FB)	– Inwieweit sehen sich Chemielehramtsstudierende in der Lage, eine klare Position bei der Bewertung des Kontextes einzunehmen? (FB)

Zur Beantwortung dieser Untersuchungsfragen wurde analog zu Studie S2 die qualitative Inhaltsanalyse nach Kuckartz (2016) eingesetzt, jedoch mit der Anpassung, dass die Überkategorien deduktiv den entwickelten konzeptionellen Ansätzen entstammten. In diesem Sinne handelte es sich um eine deduktiv-induktive qualitative Inhaltsanalyse, welche in Teilen die induktiv generierten Kategorien der Studie S2.I1

aufgriff und weiterentwickelte. Auf diese Weise konnte in Studie S3.I1&2 das so entstandene Kategoriensystem mit den konzeptionellen Überlegungen zusammengebracht werden.

Zur Beantwortung der Untersuchungsfrage eines jeden Teilbereichs wurde für jede Frage ein eigenes Kategoriensystem entwickelt.

Zur Beantwortung der Untersuchungsfrage

> Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug fachliche Aspekte? (NW-A)

des *naturwissenschaftlichen Teilbereichs* (NW) konnten sieben differente Aspekte als Kategorien ausgearbeitet werden. Ziel war es, einen Eindruck davon zu erhalten, in welchem Maße die Studierenden des Chemielehramts fachliche Inhalte als relevant bei der Bewertung erachteten.

Neben der *Effizienz* (NW-A-EF) wurden von den Studierenden beispielsweise auch die *Nachhaltigkeit* (NW-A-NA), das Verständnis für die *Funktionsweise* (NW-A-FT) von Technologien oder spezifischere Aspekte, wie die *Rohstoffherkunft* (NW-A-RH), als relevant für den Entscheidungsprozess herausgestellt (Tabelle 56).

Tabelle 56: S3.I1&2 - Kategorien der *Aspekte naturwissenschaftlicher Bewertung* (NW-A)

Überkategorie	Kategorie	Definition
Aspekte naturwissenschaftlicher Bewertung (NW-A)	NW-A-KB: *Keine naturwissenschaftliche Bewertung*	Es findet keine Bewertung auf naturwissenschaftlicher Ebene statt oder es werden nur diffuse Ansätze deutlich. Immer dann kodiert, wenn keine andere Kategorie dieser Untersuchungsfrage kodiert werden konnte.
	NW-A-EF: *Effizienz*	Die Effizienz wird als relevanter Faktor bei der Bewertung angesehen.
	NW-A-KA: *Kostenabwägung*	Die Kosten oder das Kosten-Nutzen-Verhältnis werden als relevante Faktoren bei der Bewertung angesehen.
	NW-A-NA: *Nachhaltigkeit und Auswirkungen*	Die Auswirkungen und/oder Nachhaltigkeit werden als relevante Faktoren bei der Bewertung angesehen.
	NW-A-FT: *Funktionsweise*	Die Funktionsweise wird als relevanter Faktor bei der Bewertung angesehen.
	NW-A-UM: *Umsetzbarkeit*	Die Umsetzbarkeit wird als relevanter Faktor bei der Bewertung angesehen.
	NW-A-RH: *Rohstoffherkunft*	Die Herkunft von Rohstoffen und Bestandteilen wird als relevanter Faktor bei der Bewertung angesehen.

Zur Beantwortung der Untersuchungsfrage

> Inwieweit findet bei Studierenden des Chemielehramts Fachwissen Einzug in die Bewertung des Kontextes? (NW-FB)

wurden drei Kategorien unterschieden, die verschiedene Abstufungen der Anwendung wissenschaftlichen Wissens abbildeten. Diese reichten von *nicht vorhanden* (NW-FB-NV) bis hin zu *fundiert* (NW-FB-FU) (Tabelle 57).

Tabelle 57: S3.I1&2 - Kategorien der *fachwissenschaftlichen Bewertung* (NW-FB)

Überkategorie	Kategorie	Definition	Niveau
Fachwissenschaftliche Bewertung (NW-FB)	NW-FB-NV: *Nicht vorhanden*	Es findet keine Bewertung auf Grundlage naturwissenschaftlichen Fachwissens statt.	N = 0
	NW-FB-UG: *Ungeprüft/ Unbelegt*	Zur Bewertung werden nicht überprüfte Aussagen mit naturwissenschaftlichem Inhalt herangezogen (Hörensagen).	N = 1
	NW-FB-FU: *Fundiert*	Zur Bewertung werden überprüfte Aussagen mit naturwissenschaftlichem Inhalt herangezogen.	N = 2

Angelehnt an eine skalierende Strukturierung nach Mayring (2010), welche von Kuckartz (2016) auch als evaluative qualitative Inhaltsanalyse bezeichnet wird, wurde bei der fachwissenschaftlichen Bewertung (NW-FB) eine Abstufung in drei Niveaus vorgenommen, je nachdem wie fundiert die fachliche Bewertung ausfiel.

Die Selbsteinschätzung der Studierenden hinsichtlich ihres Fachwissens zur Beantwortung der Untersuchungsfrage

> Wie schätzen Studierende des Chemielehramts ihr Fachwissen in dem gestellten Kontext ein? (NW-SF)

konnte ebenfalls in drei Kategorien kodiert werden, die von *unzureichend* (NW-SF-UZ) bis zu *umfangreich* (NW-SF-UR) reichten. Analog zur *fachlichen Bewertung* (FB) wurden diese Kategorien skalierend in Niveaus unterteilt (Tabelle 58).

Tabelle 58: S3.I1&2 - Kategorien der *Selbsteinschätzung des Fachwissens* (NW-SF)

Überkategorie	Kategorie	Definition	Niveau
Selbsteinschätzung Fachwissen (NW-SF)	NW-SF-UZ: *Unzureichend*	Das eigene Fachwissen über den Kontext wird als unzureichend bewertet.	N = 0
	NW-SF-GL: *Grundlegend*	Das eigene Fachwissen über den Kontext wird als grundlegend bewertet.	N = 1
	NW-SF-UR: *Umfangreich*	Das eigene Fachwissen über den Kontext wird als umfangreich bewertet.	N = 2

Der zweite Teilbereich der *Selbstreflexion* (SR) umfasste zwei Untersuchungsfragen. Die Untersuchungsfrage

> Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug selbstreflexive Aspekte? (SR-A)

eruierte analog zu den *Aspekten naturwissenschaftlicher Bewertung* (NW-A) die verschiedenen *Facetten der Selbstreflexion* (SR-A), wobei im Gegensatz zu den Kategorien in Studie S2.I1 nicht in *introspektive Überprüfung* (SR-IU) und das *Bewusstsein für externe und interne Einflussfaktoren* (SR-EF) differenziert wurde, da sich diese Abgrenzungen als nicht trennscharf und exakt genug herausstellten. Stattdessen wurden die *Aspekte der Selbstreflexion* (SR-A) hier in ein Bewusstsein für *interne Einflussfaktoren* (SR-A-IF) und ein Bewusstsein für *externe Einflussfaktoren* (SR-A-EF) unterschieden. *Interne Faktoren* (SR-A-IF) schließen primär selbstbezogene innere Prozesse ein, wie die eigenen *Emotionen* (SR-A-IF-EM), *Befangenheiten* (SR-A-IF-BV) oder ein Bewusstsein für *kognitive Verzerrungen* (SR-A-IF-KV), wohingegen *externe Faktoren* (SR-A-EF) äußere Einflüsse, wie die *Medien* (SR-A-EF-ME) oder *religiöse Überzeugungen* (SR-A-EF-RL), umfassen (Tabelle 59).

Tabelle 59: S3.I1&2 - Kategorien der *Aspekte der Selbstreflexion* (SR-A)

Überkategorie	Kategorie	Subkategorie	Definition
Aspekte der Selbstreflexion (SR-A)	SR-A-KS: *Keine Selbstreflexion*	Es findet keine Selbstreflexion statt oder es werden nur diffuse Ansätze deutlich. Immer dann kodiert, wenn keine andere Kategorie dieser Untersuchungsfrage kodiert werden konnte.	
	SR-A-IF: *Bewusstsein für interne Einflussfaktoren*	SR-A-IF-FW: *Fachwissen*	Das eigene Fachwissen wird bei der Bewertung reflektiert.
		SR-A-IF-PR: *Persönliche Relevanz*	Die Relevanz für sich selbst im Sinne von Vor- und Nachteilen wird bei der Bewertung reflektiert.
		SR-A-IF-MW: *Moral und Werte*	Die eigenen Werte und Moralvorstellungen werden bei der Bewertung reflektiert.
		SR-A-IF-HA: *Handlungen*	Die eigenen Handlungen werden bei der Bewertung rückwirkend betrachtet.
		SR-A-IF-EM: *Emotionen*	Die eigene Gefühlswelt wird bei der Bewertung reflektiert.
		SR-A-IF-EF: *Eigene Erfahrungen*	Die eigenen gemachten Erfahrungen werden bei der Bewertung reflektiert.
		SR-A-IF-IN: *Interesse*	Das eigene Interesse wird bei der Bewertung reflektiert.
		SR-A-IF-BV: *Befangenheit, Vorurteile*	Die eigene Befangenheit oder Vorurteile werden bei der Bewertung reflektiert.
		SR-A-IF-PK: *Persönlichkeit*	Die eigene Persönlichkeit oder Identifikation werden bei der Bewertung reflektiert.
		SR-A-IF-KV: *Kognitive Verzerrungen*	Der Einfluss kognitiver Verzerrungen wird bei der Bewertung reflektiert.
	SR-A-EF: *Bewusstsein für externe Einflussfaktoren*	SR-A-EF-SN: *Soziale Norm*	Der Einfluss der sozialen Norm wird bei der Bewertung reflektiert.
		SR-EF-EF-ME: *Medien*	Der Einfluss der Medien wird bei der Bewertung reflektiert.
		SR-EF-EF-BI: *Bildung*	Der Einfluss der Bildung wird bei der Bewertung reflektiert.
		SR-A-EF-PO: *Politik*	Der Einfluss politischer Einstellungen wird bei der Bewertung reflektiert.
		SR-A-EF-RE: *Religion*	Der Einfluss religiöser Überzeugungen wird bei der Bewertung reflektiert.

Die zweite Untersuchungsfrage zum Teilbereich der *Selbstreflexion* (SR) zielte auf das Verständnis ebenjener Begrifflichkeit ab und wurde wie folgt formuliert:

> Welches Verständnis haben Chemielehramtsstudierende von dem Begriff der Selbstreflexion? (SR-V)

Tabelle 60: S3.I1&2 - Kategorien des *Verständnisses der Selbstreflexion* (SR-V)

Überkategorie	Kategorie	Subkategorie	Definition	Niveau
Verständnis Selbstreflexion (SR-V)	SR-V-ND: *Nicht definierbar*	Selbstreflexion kann nicht definiert werden oder es liegen diffuse Definitionsansätze vor.		N = 0
	SR-V-IT: *Impliziertes Teilverständnis*	SR-V-IT-AB: *Abwägen*	Selbstreflexion wird als ein abwägender Prozess, zum Beispiel zwischen zwei Standpunkten, verstanden.	N = 1
		SR-V-IT-SV: *Selbstverbesserung*	Selbstreflexion wird als ein Prozess der Selbstverbesserung verstanden, welcher auch Selbstkritik beinhalten kann.	
		SR-V-IT-EM: *Empathie*	Selbstreflexion wird als das Hineinfühlen in andere Personen verstanden und kann auch mit dem Reflektieren der Auswirkungen des eigenen Handelns auf andere Personen einhergehen.	
		SR-V-IT-HF: *Hinterfragen*	Selbstreflexion wird als das Hinterfragen von Themen oder Dingen verstanden.	
	SR-V-ET: *Explizites Teilverständnis*	SR-V-ET-IS: *Introspektion*	Selbstreflexion wird als Zugang zu eigenen Emotionen, Gedanken oder Werten oder als Fähigkeit zur Introspektion verstanden.	N = 2
		SR-V-ET-HA: *Überdenken von Handlungen*	Selbstreflexion wird als eine ex post Bewertung des eigenen Handelns oder Verhaltens verstanden.	
		SR-V-ET-EX: *Bewusstsein für externe Prägungen*	Selbstreflexion wird als das Bewusstmachen und Reflektieren externer Prägungen verstanden.	

Die Differenzierung der Kategorien erfolgte hinsichtlich der Übereinstimmung mit der Definition der Selbstreflexion (vgl. Kap. 3.3 oder Anfang Kap. 5.3) und inwieweit Teilaspekte der Grunddefinition implizit oder explizit darin enthalten waren (Tabelle 60).

Ein *implizites Teilverständnis* (SR-V-IT) lag vor, wenn naive oder unkonkrete Definitionsansätze vorhanden waren, welche dennoch implizit Teile der Grunddefinition berührten. Stellvertretend hierfür können die Subkategorien des *Hinterfragens* (SR-V-IT-HF) oder *Abwägens* (SR-V-IT-AB) angeführt werden, welche nicht konkret mit der Definition übereinstimmen, aber Teilbereiche durch das implizierte Durchdenken oder Abwägen von Faktoren im Bewertungsprozess anschneiden. Auch die *Empathie* (SR-V-IT-EM) berührt etwa die *Introspektion* (SR-V-ET-IS) oder das *Überdenken von Handlungen* (SR-V-ET-HA) dergestalt, dass beim Hineinfühlen in eine andere Person indirekt die eigene Gefühlswelt betrachtet und reflektiert wird, insbesondere im Hinblick auf Handlungen, die möglicherweise einen Einfluss auf die entsprechende Person gehabt haben könnten.

Das *explizite Teilverständnis* (SR-V-ET) wiederum spiegelte eine informierte Vorstellung mit konkretem Bezug zu Elementen der Grunddefinition wider, wie der *Introspektion* (SR-V-ET-IS) oder dem *Bewusstsein für externe Prägungen* (SR-V-ET-EX). Daraus ergab sich die Möglichkeit einer Abstufung in Niveaus, je nach Grad des Verständnisses der Begrifflichkeit der Selbstreflexion (Tabelle 60).

Der letzte große Teilbereich, der aus dem Triadischen Modell Kritischen Denkens in der Naturwissenschaftsdidaktik entlehnt wurde, wird durch das *epistemologische Verständnis* (EP) charakterisiert und umfasst die Untersuchungsfrage

> Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug epistemologische Aspekte? (EP-A)

Zu Beantwortung dieser Frage wurden neun Kategorien definiert, wobei jede für einen anderen relevanten Faktor bei der Einschätzung der Glaubwürdigkeit einer Quelle stand. Neben der *Aktualität* (EP-A-B-AK) spielte beispielsweise auch die *Herkunft von Informationen* (EP-A-B-HK) oder die *Konsistenz mit anderen Quellen* (EP-A-B-KO) eine entscheidende Rolle bei der Einschätzung der Glaubwürdigkeit einer Quelle (Tabelle 61).

Tabelle 61: S3.I1&2 - Kategorien der *Aspekte des epistemologischen Verständnisses* (EP-A)

Überkategorie	Kategorie	Subkategorie	Definition
Aspekte des epistemologischen Verständnisses (EP-A)	EP-A-KB: *Keine Bewertung mittels epistemologischer Aspekte*	Erfüllt, wenn keine Bewertung von Quellen mittels epistemologischer Aspekte vorgenommen wird. Immer dann kodiert, wenn keine andere Kategorie der Untersuchungsfrage kodiert werden konnte.	
	EP-A-B: *Bewertung mittels epistemologischer Aspekte*	EP-A-B-EX: *Expertise*	Die Autorität und Expertise der*s Autor*in werden als relevante Faktoren bei der Einschätzung der Vertrauenswürdigkeit einer Quelle angesehen.
		EP-A-B-AK: *Aktualität*	Die Aktualität einer Information wird als relevanter Faktor bei der Einschätzung der Vertrauenswürdigkeit einer Quelle angesehen.
		EP-A-B-IK: *Interessenkonflikte*	Interessenkonflikte oder Subjektivität werden als relevante Faktoren bei der Einschätzung der Vertrauenswürdigkeit einer Quelle angesehen.
		EP-A-B-PR: *Präsentation*	Die Art der Präsentation wird als relevanter Faktor bei der Einschätzung der Vertrauenswürdigkeit einer Quelle angesehen.
		EP-A-B-HK: *Herkunft*	Die Herkunft einer Information wird als relevanter Faktor bei der Einschätzung der Vertrauenswürdigkeit einer Quelle angesehen.
		EP-A-B-ME: *Methodik*	Die verwendeten Methoden oder das Forschungsdesign werden als relevante Faktoren bei der Einschätzung der Vertrauenswürdigkeit einer Quelle angesehen.
		EP-A-B-KO: *Konsistenzprüfung*	Die Übereinstimmung einer Information mit anderen Quellen wird als relevanter Faktor bei der Einschätzung der Vertrauenswürdigkeit einer Quelle angesehen.
		EP-A-B-TR: *Transparenz*	Die offene und transparente Angabe von Herkunftsinformationen wird als relevanter Faktor bei der Einschätzung der Vertrauenswürdigkeit einer Quelle angesehen.
		EP-A-B-SE: *Soziokulturelle Eingebundenheit*	Die Berücksichtigung soziokultureller Einflüsse wird als relevanter Faktor bei der Einschätzung der Vertrauenswürdigkeit einer Quelle angesehen.

Neben den dem Triadischen Modell Kritischen Denkens entlehnten Teilbereichen standen zwei weitere Bereiche im Erkenntnisinteresse. Angelehnt an die Kategorie zum *Verständnis der Begrifflichkeit Kritischen Denkens* (KD-V) aus Studie S2 verfolgte auch die Studie S3 die Beantwortung der Untersuchungsfrage

> Welches Verständnis haben Chemielehramtsstudierende von der Formulierung kritisch denken oder kritisch sein? (KD-V)

mit der Kodierung zugehöriger Aussagen. Mithilfe des neuen Datenmaterials konnte auf diese Weise die in Studie S2 gelegte Grundlage weiter ausgeschärft und differenziert werden (Tabelle 62).

Tabelle 62: S3.I1&2 - Kategorien des *Verständnisses von Kritischem Denken* (KD-V)

Überkategorie	Kategorie	Subkategorie	Definition
Verständnis von Kritischem Denken (KD-V)	KD-V-NA: *Naive Vorstellung*	KD-V-NA-ND: *Nicht definierbar*	Kritisches Denken kann nicht definiert werden oder es liegt eine diffuse Definition vor.
		KD-V-NA-HF: *Hinterfragen*	Kritisches Denken wird als das Hinterfragen von Dingen oder Themen verstanden, wobei dies nicht weiter ausgeführt wird.
		KD-V-NA-SK: *Skepsis*	Kritisches Denken wird als grundsätzlich skeptische und ungläubige Haltung verstanden.
		KD-V-NA-SV: *Selbstverbesserung/ -kritik*	Kritisches Denken wird als ein Prozess der Selbstverbesserung verstanden, welcher aus tadelnder Selbstkritik hervorgeht.
		KD-V-NA-TK: *Tadeln/Kritisieren*	Kritisches Denken wird als das Kritisieren oder Tadeln anderer verstanden.
		KD-V-NA-SD: *Selbstständiges Denken*	Kritisches Denken wird als eigenständiges Denken definiert, wobei die Dimension nicht weiter ausgeführt wird.
	KD-V-TG: *Teilbereich der Grunddefinition*	KD-V-TG-SR: *Selbstreflexion*	Kritisches Denken wird als das Hinterfragen der eigenen Person oder den eigenen Prägungen im Sinne einer Selbstreflexion verstanden.
		KD-V-TG-WI: *Wissen*	Kritisches Denken wird als das Hinterfragen und Hinzuziehen von externen Informationen oder eigenem Wissen verstanden.
		KD-V-TG-EP: *Epistemologisches Verständnis*	Kritisches Denken wird als das Hinterfragen und Bewerten von Quellen und ihrer Glaubwürdigkeit verstanden.
	KD-V-MT: *Zusammenspiel mehrerer Teilbereiche der Grunddefinition*	Kritisches Denken wird als ein (multiperspektivisches) Zusammenspiel mindestens zweier Teilbereiche von Selbstreflexion, Wissen und epistemologischem Verständnis verstanden.	

Es ließen sich, vergleichbar mit der Strukturierung bei Kategorie des *Verständnisses der Selbstreflexion* (SR-V), drei Niveaustufen unterscheiden. Niveau 0 wurde kodiert, wenn die Vorstellungen von dem Begriff des Kritischen Denkens diffus oder naiv definiert vorlagen, wie die unkonkrete Benennung als *Hinterfragen* (KD-V-NA-HF) oder das Verständnis als *Skepsis* (KD-V-NA-SK). Wurden konkrete Teilbereiche der Grunddefinition oder des Triadischen Modells Kritischen Denkens benannt, wurde die Kategorie *Teilbereich der Grunddefinition* (KD-V-TG) als Niveau 1 kodiert, welche

abermals in die drei Teilbereiche *Selbstreflexion* (KD-V-TG-SR), *Wissen* (KD-V-TG-WI) und *epistemologisches Verständnis* (KD-V-TG-EP) differenziert wurde. Das höchste Niveau 2 wurde dann erreicht, wenn nicht nur ein Teilbereich als Bestandteil Kritischen Denkens benannt wurde, sondern das *Zusammenspiel mindestens zweier Teilbereiche der Grunddefinition* (KD-V-MT) deutlich wurde. Die so entstandene Kategorie stellt damit den multiperspektivischen Charakter Kritischen Denkens heraus.

Die letzte Untersuchungsfrage

> Inwieweit sehen sich die Studierenden des Chemielehramts in der Lage, eine klare Position bei der Bewertung des Kontextes einzunehmen? (FB)

kann als Zusatzfragestellung bezeichnet werden, um gemeinsam mit den Antworten der Untersuchungsfragen zur *Selbsteinschätzung des Fachwissens* (SF) und der *fachwissenschaftlichen Bewertung* (NW-FB) erweiterte kombinierte Schlüsse zuzulassen. Diese drei Fragestellungen erwiesen sich erst in ihrer Kombination als gewinnbringend, da sie es ermöglichten, irrationale Urteile besser herauszustellen, wie dies bereits in Studie S2.I1 erprobt wurde. Die eigene *Fähigkeit zur Bewertung* (FB) wurde seitens der Teilnehmenden derart eingeschätzt, dass sie sich entweder *nicht in der Lage* (FB-NL), *unter Einschränkungen in der Lage* (FB-UE) oder *in der Lage* (FB-IL) sahen, sich im gegebenen Kontext zu positionieren. Dementsprechend wurde auch eine Strukturierung im Sinne einer Niveauabstufung eingesetzt (Tabelle 63).

Tabelle 63: S3.I1&2 - Kategorien der *Fähigkeit zur Bewertung* (FB)

Überkategorie	Kategorie	Definition	Niveau
Fähigkeit zur Bewertung (FB)	FB-NL: *Sieht sich nicht in der Lage*	Es wird sich nicht zugetraut, eine Entscheidung treffen zu können.	N = 0
	FB-UE: *Sieht sich unter Einschränkungen in der Lage*	Es wird sich mit Einschränkungen oder Vorbehalt zugetraut, eine Entscheidung treffen zu können. Entscheidungstendenzen liegen vor.	N = 1
	FB-IL: *Sieht sich in der Lage*	Es wird sich zugetraut, eine Entscheidung treffen zu können.	N = 2

Die ausführlichen Kodierleitfäden zu den einzelnen Kategorien der Interviewstudie S3.I1&2 sind dem Anhang (A.3.1.1) beigefügt.

Nach der Kodierung der Transkripte mit MAXQDA auf Grundlage der Kodierleitfäden wurden die Anzahlen der vergebenen Kategorien in Excel tabellarisch aufgetragen und sowohl die Summen pro Kategorie als auch pro Proband*in dargestellt.

In diesem Zuge wurden auch die Pre-Interviews der Befragten aus Studie S2.I1 auf Basis der überarbeiteten Kodierleitfäden rekodiert. Zusätzlich wurden, wie zuvor schon in Studie S1 und S2, die Einzelzählungen eingesetzt, da es weniger von Interesse war, wie häufig ein bestimmter Aspekt von einer*m Teilnehmenden benannt wurde, sondern wie viele verschiedene Aspekte im Sinne der Multiperspektivität im Bewertungsprozess eine Rolle spielten. Da die Aspekte der Bewertung (NW-A, SR-A und EP-A) sowie das Verständnis der Begrifflichkeiten der Selbstreflexion und des Kritischen Denkens (SR-V und KD-V) von besonderem Interesse waren, wurde auf Grundlage dieser Werte die Maße der zentralen Tendenz und die Streumaße berechnet und die Veränderungen zwischen Pre- und Post-Test grafisch mithilfe von Netzdiagrammen veranschaulicht. Zudem wurde mithilfe des Wilcoxon-Tests ein Signifikanztest durchgeführt, um zu ermitteln, ob signifikante Unterschiede zwischen Pre- und Post-Test bestanden. Nähere Erläuterungen zu den verwendeten Instrumenten der deskriptiven Statistik und der Signifikanztests sind im Kapitel 4.3.2 aufgeführt.

5.3.1.2 Ergebnisse

Bei der Betrachtung der zur Bewertung hinzugezogenen *naturwissenschaftlichen Aspekte* (NW-A) benannten die Befragten im Pre-Test durchschnittlich 2,6 verschiedene Aspekte als für sie relevant (Tabelle 64). 70 % (7 von 10) der Teilnehmenden benannten die *Nachhaltigkeit oder Auswirkungen auf die Natur* (NW-A-NA) als wichtiges Entscheidungskriterium, was in Fragen, wie „[W]as ist jetzt besser für die Umwelt?“ (S3.I1-SS23-B-B3_5-XJ19MF, Pos. 15), deutlich wurde. Die zweithäufigste Kategorie war die *Kostenabwägung* (NW-A-KA), die sechs von zehn Personen (60 %) im Bewertungsprozess berücksichtigten. Der dritte große Faktor, welcher ebenfalls von 60 % der Teilnehmenden in den Bewertungsprozess inkludiert wurde, war die tatsächliche *Effizienz* (NW-A-EF) der Technologie. Aussagen, wie „Ich würde mich erst mal erkundigen, worin die stärkste Effizienz ist [sic]“ (S2.I1-Wi22_23-A1-B2_3-XJ36JE, Pos. 26), können stellvertretend für Nennungen und Fragen dieser Kategorie angeführt werden. Weit weniger relevant war für die Studierenden die tatsächliche *Funktionsweise* (NW-A-FT) der Technik. Fragen in dieser Richtung wurden kaum gestellt. Nur ein*e Proband*in stellte Fragen, wie „Was gibt es da eigentlich für Verfahren?“ (S2.I1-Wi22_23-B-B2_3-XF21FM, Pos. 80), welche ein besseres Verständnis der Technologie selbst anstrebten. Obwohl die *Auswirkungen und die Nachhaltigkeit* (NW-A-NA) vom Großteil der Befragten als relevanter Faktor erachtet wurden, stellten nur zwei Teilnehmende die Frage nach der *Herkunft der Rohstoffe* (NW-A-RH) zur Fertigung der Technologien. Gerade der Abbau dieser macht aber einen nicht unerheblichen Teil der damit verknüpften Umweltrisiken aus, besonders, wenn man auf

die multifaktoriellen gesellschaftlichen, umweltbezogenen, politischen oder sozialen Probleme schaut, die beispielsweise mit dem Abbau Seltener Erden einhergehen.

Tabelle 64: S3.I1&2 - Ergebnisse der Interviewerhebung, *naturwissenschaftliche Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (NW), *N* = 10

Kat.	1. UK	2. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
NW: Naturwissenschaftliches Wissen	NW-A: Aspekte der naturwissenschaftlichen Bewertung	NW-A-KB: Keine Bewertung	0	0	0	0
		NW-A-EF: Effizienz	6	6	10	11
		NW-A-KA: Kostenabwägung	6	5	12	11
		NW-A-NA: Auswirkungen & Nachhaltigkeit	7	9	14	25
		NW-A-FT: Funktionsweise	1	7	2	12
		NW-A-UM: Umsetzbarkeit	3	7	7	18
		NW-A-RH: Rohstoffherkunft	2	5	3	9
		NW-A-NO: Notwendigkeit	1	2	1	2
Summe NW-A					**49**	**85**
Einzelzählung NW-A			**26**	**41**		

Im Gegensatz zum Pre-Test verbesserte sich die Bewertungsgrundlage in den durchgeführten Post-Tests deutlich und es wurden im Schnitt 4,1 differente Perspektiven bei der Bewertung hinzugezogen. *Auswirkungen und Nachhaltigkeit* (NW-A-NA) blieb die häufigste benannte Kategorie, welche von neun der zehn Befragten (90 %) berücksichtigt wurde. Während die *Effizienz* (NW-A-EF) weiterhin von 60 % (6 von 10) der Teilnehmenden berücksichtigt wurde, stieg der Anteil der Personen, die ein Verständnis der Funktionsweise einforderten, auf 70 % (7 von 10) und Verständnisfragen, wie „[W]ie sind denn die Herstellungsprozesse für Solar- oder Windenergie?“ (S3.I2-SS23-A-B3_5_EF28SH, Pos. 13), zeigten eine deutliche Häufung. Auch Fragen, wie „[W]ie weit ist das machbar in der Zukunft?“ (S3.I2-SS23-B-B2_3-SV32BT, Pos. 19), traten vermehrt auf und fielen unter die Kategorie der *Umsetzbarkeit* (NW-A-UM). Der Anteil der Teilnehmenden, die diese Kategorie nannten, stieg von ursprünglich 30 % auf 70 %. Reflektierten im Pre-Test nur zwei Personen die *Herkunft der Rohstoffe* (NW-A-RH), so waren es im Post-Test bereits 50 % der Befragten (5 von 10). In fast allen Kategorien konnten deutliche Anstiege der Berücksichtigung durch die Teilnehmenden verzeichnet werden, welche sich fachlich im Post-Interview breiter aufstellten als noch zu Beginn im Pre-Test. Insgesamt zeigten 8 Teilnehmende (80 %) erweiterte Bewertungsperspektiven im Vergleich zum Pre-Test.

Die Ergebnisse der *selbstreflexiven Aspekte* (SR-A) bei der Bewertung des Kontextes zeigen im Pre-Test einen Mittelwert von 2,5 verschiedenen *internen Faktoren* (SR-A-IF) und 1,7 unterschiedlichen *externen Faktoren* (SR-A-EF) pro Proband*in (Tabelle 65).

Tabelle 65: S3.I1&2 - Ergebnisse der Interviewerhebung, *selbstreflexive Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (NW), *N* = 10

Kat.	1. UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
SR: Selbstreflexion	SR-A: Aspekte der Selbstreflexion	SR-A-KS: Keine Selbstreflexion		0	0	0	0
		SR-A-IF: Interne Faktoren & introspektive Überprüfung	SR-A-IF-FW: Fachwissen	3	8	4	17
			SR-A-IF-PR: Persönliche Relevanz	6	8	21	21
			SR-A-IF-MW: Moral und Werte	3	6	4	12
			SR-A-IF-HA: Handlungen	1	4	1	6
			SR-A-IF-EM: Emotionen	1	4	1	11
			SR-A-IF-EF: Erfahrungen	3	4	4	5
			SR-A-IF-IN: Interesse	2	5	4	6
			SR-A-IF-BV: Befangenheit, Vorurteile	6	10	9	26
			SR-A-IF-PK: Persönlichkeit	0	0	0	0
			SR-A-IF-KV: Kognitive Verzerrungen	0	5	0	10
		Summe SR-A-IF				**48**	**114**
		Einzelzählung SR-A-IF		**25**	**54**		
		SR-A-EF: Bewusstsein für externe und interne Einflussfaktoren	SR-A-EF-SN: Soziale Norm/Umfeld	9	8	24	36
			SR-A-EF-ME: Medien	5	5	9	8
			SR-A-EF-BI: Bildung	1	3	2	5
			SR-A-EF-PO: Politik	2	3	4	3
			SR-A-EF-RE: Religion	0	0	0	0
		Summe SR-A-EF				**39**	**52**
		Einzelzählung SR-A-EF		**17**	**19**		

Die am häufigsten reflektierten internen Aspekte waren die *persönliche Relevanz* (SR-A-IF-PR) für sich selbst und die potenzielle *persönliche Befangenheit* (SR-A-IF-BV), welche von jeweils 60 % (6 von 10) der Befragten angeführt wurden. Die Aussage „Also ich würd [sic] sagen, die erste Frage wäre so, was wäre für mich am besten […]" (S2.I1-Wi22_23-A2-B2_3-XJ27BO, Pos. 32) verdeutlicht dabei die hohe Priorität, welche die *persönliche Relevanz* (SR-A-IF-PR) im Entscheidungsprozess einnehmen kann, bevor andere Faktoren oder gar fachliche Fragestellungen in Betracht gezogen werden. Deutlich weniger wurden die anderen internen Faktoren, wie *Moral und Werte* (SR-A-IF-MW) oder gemachte *Erfahrungen* (SR-A-IF-EF), reflektiert und auch die Reflexion des *eigenen Wissensstandes* (SR-A-IF-FW) wurde nur von drei der zehn Befragten (30 %) berücksichtigt.

Als *externe Einflussfaktoren* (SR-A-EF) reflektierten 90 % der Befragten (9 von 10) den Einfluss ihrer *sozialen Norm* (SR-A-EF-SN), welche in den Interviews größtenteils als der stärkste Einflussfaktor benannt wurde. Dies fasste ein*e Proband*in wie folgt zusammen: „Ich würde sagen, am stärksten ist tatsächlich der Einfluss von Freunden, Familie, also Personen, die mir nahestehen" (S3.I1-SS23-B-B3_5-EF28SH, Pos. 45). Die Hälfte der befragten Personen reflektierte zudem den Einfluss der *Medien* (SR-A-EF-ME) auf ihre Meinung.

In der Post-Erhebung wurden im Schnitt 1,9 differente *externe Einflussfaktoren* (SR-A-EF) reflektiert, was nur einen kleinen Anstieg im Vergleich zur Pre-Erhebung darstellt. Dennoch wurde auch hier eine vermehrte Benennung der *politischen Einflüsse* (SR-A-EF-PO) und der eigenen *Bildung* (SR-A-EF-BI) deutlich, die von jeweils 30 % der Befragten (3 von 10) angeführt wurden. Die größte Veränderung war bei den *internen Faktoren* (SR-A-IF) zu verzeichnen, die im Mittel von 2,5 auf 5,4 verschiedene reflektierte Aspekte pro Proband*in stieg. Der am häufigsten und voll allen Teilnehmenden im Post-Interview berücksichtigte Aspekt war die *persönliche Befangenheit* (SR-A-IF-BV) und der Einfluss, den diese auf die Meinungsbildung nehmen kann. Dies wird beispielsweise in der Aussage „Aber ein Faktor ist sicherlich meine jetzige Meinung, die ich da schon drüber hab" (S3.I2-SS23-A-B3_5_WJ32FE, Pos. 23) deutlich. 80 % der Befragten (8 von 10) reflektierten die *persönliche Relevanz* (SR-A-IF-PR) und der Anteil an Teilnehmenden, die ihr *Fachwissen* (SR-A-IF-FW) in der Bewertung berücksichtigten, stieg ebenfalls auf 80 %. Die Reflexion von *Moral und Werten* (SR-A-IF-MW) in Fragen, wie „inwieweit kann man das noch moralisch auch unterstützen?" (S2.I2-SS23-B-B3_5_XJ36JE, Pos. 21), fand bei fast zwei Dritteln der interviewten Personen statt, knapp ein Drittel mehr als noch zuvor im Pre-Test. Das Bewusstsein für *kognitive Verzerrungen* (SR-A-IF-KV) wurde im Post-Test ebenfalls bei

50 % der Befragten (5 von 10) deutlich, was im Gegensatz zum Pre-Test, in dem kognitive Verzerrungen von keiner Person thematisiert wurden, eine Erhöhung darstellt. Insgesamt konnten bei allen Teilnehmenden Verbesserungen in Form von erweiterten Perspektiven bei der Bewertung im Gegensatz zum Pre-Interview festgestellt werden, wobei der Schwerpunkt auf dem Bewusstsein für *interne Faktoren* (SR-A-IF) lag. Doch auch im Bereich der *externen Faktoren* (SR-A-EX) waren bei 50 % der Befragten erweiterte Perspektiven zu verzeichnen.

Der Bereich des *epistemologischen Verständnisses* (EP) stellt den dritten Teilbereich Kritischen Denkens dar (Tabelle 66). Im Pre-Test fanden sich bei 50 % der Befragten (5 von 10) *keine Teilaspekte* (EP-A-KB) für die Bewertung von Informationsquellen oder zur Einschätzung von Informationen. Nur eine Person traf Aussagen, welche die *Herkunft* (EP-A-B-HK) oder Veröffentlichungsplattform einer Information als relevant hervorhoben. Jeweils zwei Personen sahen die *Expertise* (EP-A-B-EX) der*s Autor*in oder die *Übereinstimmung mit anderen Quellen* (EP-A-B-KO) als wichtig bei der Bewertung von Quellen an. Die Anzahl differenter Kriterien zur Bewertung von Informationen pro Proband*in lag bei 0,5.

Tabelle 66: S3.I1&2 - Ergebnisse der Interviewerhebung, *epistemologische Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (NW), $N = 10$

Kat.	1. UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
EP: Epistemologie	EP-A: Aspekte der Epistemologie	EP-A-KB: Keine Bewertung der Quelle		5	0	5	0
		EP-A-B: Bewertung auf Basis epistemologischer Aspekte	EP-A-B-EX: Expertise	2	5	3	9
			EP-A-B-AK: Aktualität	0	1	0	1
			EP-A-B-IK: Interessenskonflikte	0	6	0	10
			EP-A-B-PR: Präsentation	0	3	0	5
			EP-A-B-HK: Herkunft	1	10	2	20
			EP-A-B-ME: Methodik	0	3	0	3
			EP-A-B-KO: Konsistenzprüfung	2	5	2	5
			EP-A-B-TR: Transparenz	0	3	0	3
			EP-A-B-SE: Soziokulturelle Eingebundenheit	0	2	0	2
Summe EP-A-B						**7**	**58**
Einzelzählung EP-A-B				**5**	**38**		

Dieser Mittelwert stieg im Post-Interview erheblich auf 3,8 verschiedene Aspekte pro Proband*in. Alle Teilnehmenden stellten Fragen, wie „Woher kommen diese Informationen? Was sind das für Quellen?“ (S3.I2-SS23-A-B3_5_EF28SH, Pos. 13), welche die Relevanz der *Herkunft* (EP-A-B-HK) oder Veröffentlichungsplattform für die Teilnehmenden widerspiegelten. 60 % (6 von 10) gaben potenzielle *Interessenkonflikte* (EP-A-B-IK) als Einfluss auf die Glaubwürdigkeit von Informationen und Quellen an, was sich in Fragen, wie „Wer steckt vielleicht dahinter? Was könnten die für Interessen haben?“ (S3.I2-SS23-A-B3_5_EF28SH, Pos. 13), niederschlug. 50 % der Befragten (5 von 10) empfanden die *Expertise der*s Autor*in* (EP-A-B-EX) und das *Übereinstimmen von Informationen mit anderen Quellen* (EP-A-B-KO) als wichtig zur Einschätzung der Glaubwürdigkeit. Drei von zehn Teilnehmenden (30 %) gaben zudem an, die *Transparenz von Quellen* (EP-A-B-TR) im Bewertungsprozess als wichtig zu erachten, und sogar zwei Personen empfanden die Berücksichtigung der *soziokulturellen Eingebundenheit* (EP-A-B-SK) von Quellen als notwendig. In Summe erweiterten sich bei neun der zehn befragten Personen die bei der Bewertung hinzugezogenen Perspektiven deutlich.

Bei der Auswertung der Antworten auf die Frage nach dem *Verständnis der Selbstreflexion* (SR-V) erreichten sechs Teilnehmende bereits im Pre-Test das Niveau 2 und damit ein *explizites Teilverständnis* (SR-V-ET) (Tabelle 67). Das hier auftretende Verständnis charakterisierte die Selbstreflexion größtenteils als die *ex post Betrachtung eigener Handlungen* (SR-V-ET-HA) und nur drei Befragte verstanden darunter die *Fähigkeit zur Introspektion* (SR-V-ET-IS). Ein *Bewusstsein für prägende externe Faktoren* (SR-V-ET-EX) als Teil der Selbstreflexion lag zu diesem Zeitpunkt bei keiner Person vor. Häufiger traten naive Vorstellungen mit *implizitem Teilverständnis* (SR-V-IT) auf, wie die Interpretation von Selbstreflexion als selbstverbessernder Prozess, welcher in den meisten Fällen mit tadelnder Selbstkritik einhergeht. Dies wird beispielsweise durch Aussagen, wie „[I]ch bin auch selbstkritisch und würde sagen, ich erkenne auch fehlerhafte Verhaltensweisen“ (S3.I1-SS23-B-B3_5-XJ19MF, Pos. 71), hervorgehoben. Lediglich eine Person verblieb auf Niveau 0.

Tabelle 67: S3.I1&2 - Ergebnisse der Interviewerhebung, *Verständnis Selbstreflexion* (SR-V), *N* = 10

Kat.	1. UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
SR: Selbstreflexion	SR-V: Verständnis Selbstreflexion	SR-V-ND: nicht definierbar		1	0	1	0	N=0
		SR-V-IT: Implizites Teilverständnis	SR-V-IT-AB: Abwägen	0	0	0	0	N=1
			SR-V-IT-SV: Selbstverbesserung /Kritik	3	0	6	0	
			SR-V-IT-EM: Empathie	2	1	2	2	
			SR-V-IT-HF: Hinterfragen	2	1	2	1	
		Summe SR-V-IT				**10**	**3**	
		Einzelzählung SR-V-IT		**7**	**2**			
		SR-V-ET: Explizites Teilverständnis	SR-V-ET-IS: Introspektion	3	8	6	16	N=2
			SR-V-ET-HA: Überdenken von Handlungen	6	6	11	9	
			SR-V-ET-EX: Bewusstsein für externe Prägungen	0	4	0	6	
		Summe SR-V-ET				**17**	**31**	
		Einzelzählung SR-V-ET		**9**	**18**			

Im Post-Test hingegen erreichten alle Befragten das Niveau 2 und damit ein *explizites Teilverständnis* (SR-V-ET). 80 % der interviewten Personen (8 von 10) fassten unter dem Begriff der Selbstreflexion im Post-Test die *Introspektion in die eigene Gedanken- oder Gefühlswelt* (SR-V-ET-IS), was eine Person wie folgt beschrieb:

> Das ist im Prinzip in meinem Verständnis so ein bisschen dieses, diese Metaebene, also das Denken über das Denken. Oder das Denken über die eigenen Emotionen. Also sich mit den eigenen Denkprozessen oder den eigenen Emotionen halt auf einer Ebene beschäftigen, die so ein bisschen pseudoobjektiv von oben da draufschaut (S3.I2-SS23-A-B3_5_EF28FH, Pos. 31).

60 % (6 von 10) gaben zudem an, das *Überdenken von Handlungen* (SR-V-ET-HA), in Fragen wie „Warum habe ich das gemacht? Warum habe ich das gesagt oder so gehandelt?“ (S3.I2-SS23-A-B2_3-XJ19MF, Pos. 29), zur Selbstreflexion zu zählen. Ergänzend dazu sahen vier von zehn Befragten auch das *Bewusstsein für externe Prägungen* (SR-V-ET-EX) als Teil der Selbstreflexion an. Naivere Ansichten verschwanden fast gänzlich. Zusammengefasst verbesserte sich bei 70 % der Teilnehmenden (7 von 10) das Verständnis bezüglich des Begriffs der Selbstreflexion. Von den drei verbleiben-

den Personen wiesen zwei bereits im Pre-Test ein informiertes Verständnis der Begrifflichkeit auf, wodurch in diesen Fällen keine Veränderung festgestellt werden konnte.

Im Gegensatz zum *Verständnis der Selbstreflexion* (SR-V) erreichten im Pre-Test nur zwei der zehn interviewten Personen (20 %) beim *Verständnis des Kritischen Denkens* (KD-V) das Niveau 1 und keine Person Niveau 2 (Tabelle 68). Die Vorstellungen vom Begriff des Kritischen Denkens stellten sich bei den Studierenden als stark subjektiv geladen und schwer greifbar dar. Fünf der zehn Teilnehmenden (50 %) verstanden unter *kritisch denken* eine grundlegend ungläubige und *skeptische Haltung* (KD-V-NA-SK), welche sich in Äußerungen, wie „Kritisch denken bedeutet für mich irgendwie auch schwer zu überzeugen sein" (S3.I1-SS23-A-B3_5-SJ16DI, Pos. 64), ausdrückte. Die Aussage „Kritisch sein und kritisch denken ist für mich eigentlich hinterfragen und nicht alles so hinnehmen, wie es ist" (S3.I1-SS23-A-B3_5-RU21AS, Pos. 71) beinhaltet neben dem Verständnis als Skepsis außerdem die Kategorie des *Hinterfragens* (KD-V-NA-HF), welche bei 40 % der Befragten (4 von 10) kodiert wurde. Diese Äußerungen verblieben oft oberflächlich und unkonkret, sodass unklar blieb, was oder wie genau hinterfragt werden sollte. Nur zwei Personen streiften *Teilbereiche der Grunddefinition* (KD-V-TG) im Bereich des *Wissens* (KD-V-TG-EI) und des *epistemologischen Verständnisses* (KD-V-TG-EP).

Im Post-Test zeigten sich dagegen deutlich informiertere Vorstellungen von der Begrifflichkeit des Kritischen Denkens. Alle Teilnehmenden erreichten im Post-Test mindestens das Niveau 1, sieben Befragte sogar Niveau 2, indem *mehr als ein Teilbereich der Grunddefinition* (KD-V-MT) genannt wurde. Der Teilbereich des *Wissens* (KD-V-TG-WI) wurde von 90 % der interviewten Personen (9 von 10) als relevanter Bestandteil Kritischen Denkens angeführt. So bedeutete Kritisches Denken unter anderem für eine Person nicht nur das Abrufen von Faktenwissen, sondern „dass man sich ein gewisses Wissen angeeignet hat, mit dem man jetzt […] zu einer begründeten Stellungnahme kommt, bei der man Pro und Contra mit reinbringen kann" (S2.I2-SS23-B-B3_5-XJ27BO, Pos. 47). 80 % der Teilnehmenden (8 von 10) ordneten die Fähigkeit zur *Selbstreflexion* (KD-V-TG-SR) dem Kritischen Denken zu. Eine interviewte Person umschrieb diese mit folgenden Worten:

> Also je mehr man bewusst ist über das, was, wer man ist und wie man ist und warum man so ist und was einen beeinflusst und was eben auch nicht vielleicht, desto besser kann man auf viele Sachen auch reagieren. Und das ist dieses, was ich meine mit kritisch sein, also nicht nur Dingen und anderen Menschen gegenüber, sondern aus sich selbst (S2.I2-SS23-A-B2_3-BN10LB, Pos. 41).

Tabelle 68: S3.I1&2 - Ergebnisse der Interviewerhebung, *Verständnis Kritisches Denken* (KD-V), *N* = 10

Kat.	1. UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
KD: Kritisches Denken	KD-V: Verständnis von KD	KD-V-NA: naive Vorstellung von KD	KD-V-NA-ND: nicht definierbar	2	0	2	0	N=0
			KD-V-NA-HF: Hinterfragen	4	2	7	2	
			KD-V-NA-SK: Skepsis	5	3	5	5	
			KD-V-NA-SV: Selbstverbesserung	2	0	2	0	
			KD-V-NA-TK: Tadeln/Korrigieren	2	0	3	0	
			KD-V-NA-SD: Selbstständiges Denken	1	2	1	2	
		Summe KD-V-NA:				**20**	**9**	
		Einzelzählung KD-V-NA:		**16**	**7**			
		KD-V-TG: Teilbereich der Grund-definition	KD-V-TG-SR: Selbstreflexion	0	8	0	18	N=1
			KD-V-TG-WI: Wissen	1	9	2	13	
			KD-V-TG-EP: Epistemologisches Verständnis	1	6	1	10	
		Summe KD-V-TG:				**3**	**41**	
		Einzelzählung KD-V-TG:		**2**	**23**			
		KD-V-MT:		**0**	**7**	**0**	**7**	**N=2**

Auch das *epistemologische Verständnis* (KD-V-TG-EP) fand bei 60 % der Teilnehmenden (6 von 10) Einzug in die Definitionen Kritischen Denkens. Als Beispiel kann folgendes Zitat angeführt werden:

> Auch wenn ich irgendwelche Statistiken sehe, dass ich halt gucke, woher kommt das? Wer hat das gemacht? Wann wurde das gemacht? Also die Zeit, in welcher Zeit das gemacht wurde, spielt ja auch eine Rolle. Der Ort halt irgendwo auch, ob es jetzt USA ist, Deutschland ist auch nochmal ein Unterschied von der Gesellschaft (S2.I2-SS23-A-B2_3-BN10LB, Pos. 51).

Naivere Definitionsansätze traten deutlich vermindert auf und nur noch in den Bereichen *Skepsis* (KD-V-NA-SK), *Hinterfragen* (KD-V-NA-HF) und *selbstständiges Denken* (KD-V-NA-SD). Insgesamt zeigte sich bei allen Teilnehmenden eine deutliche Verbesserung im Verständnis von Kritischem Denken, die sich daran zeigte, dass

alle Teilnehmenden mindestens die Niveaustufe 1 erreichten. Sieben Personen erreichten überdies Niveaustufe 2 und verstanden Kritisches Denken als das Zusammenspiel mehrerer Teilbereiche.

Die letzten drei Untersuchungsfragen bezüglich der *Selbsteinschätzung des Fachwissens* (NW-SF), der *fachwissenschaftlichen Bewertung* (NW-FB) und der *Fähigkeit zur Bewertung* (FB), die als ergänzende Fragestellungen zum Hauptuntersuchungsinteresse formuliert wurden, entfalten ihre größte Aussagekraft in der verknüpften Betrachtung der Ergebnisse. Im Pre-Interview sah nur eine Person ihr Fachwissen als *umfangreich* (NW-SF-UR) an, konnte aber keine *fundierte* (NW-FB-FU) fachliche Position im Interview beziehen. Generell sahen sich die Studierenden nicht in der Lage, zu den gegebenen Kontexten eine fachliche Position einzunehmen, die über ein Hörensagen und damit Niveaustufe 1 der *fachlichen Bewertung* (NW-FB) hinausging. Drei Teilnehmende schätzten ihr Wissen zum gegebenen Kontext als *grundlegend* (NW-SF-GL) ein, drei Personen verblieben ihrer eigenen Einschätzung nach auf Niveau 0 (NW-SF-UZ). Wiederum drei Befragte machten widersprüchliche Angaben, was zu Kodierungen des Fachwissens sowohl als *unzureichend* (NW-SF-UZ) als auch als *grundlegend* (NW-SF-GL) führte (Tabelle 69). Im Vergleich dazu zeigte sich die *Selbsteinschätzung des Fachwissens* (NW-SF) im Post-Test deutlich defensiver und der Großteil der Teilnehmenden (7 von 10, 70 %) verblieb bei der Einschätzung des Wissens als *unzureichend* (NW-SF-UZ).

Tabelle 69: S3.I1&2 - Ergebnisse der Interviewerhebung, *Selbsteinschätzung Fachwissen (NW-SF) und Fachwissenschaftliche Bewertung (NW-FB*), *N* = 10

Kat.	1. UK	2. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
NW: Naturwissenschaf	NW-SF: Selbsteinschätzung Fachwissen	NW-SF-UZ: Unzureichend	6	7	7	15	**N=0**
		NW-SF-GL: Grundlegend	6	1	8	1	**N=1**
		NW-SF-UR: Umfangreich	1	0	1	0	**N=2**
NW: Naturwissenschaftliche	NW-FB: Fachwissenschaftliche Bewertung	NW-FB-NV: Nicht vorhanden	2	3	2	3	**N=0**
		NW-FB-UG: Ungeprüft/ Unbelegt	8	7	13	14	**N=1**
		NW-FB-FU: Fundiert	0	0	0	0	**N=2**

Von den zehn Studierenden sahen sich im Pre-Test sechs (60 %) *nicht in der Lage* (FB-NL) eine klare Position zu beziehen, wohingegen zwei sich zumindest *unter Einschränkungen* (FB-UE) und zwei deutlich *fähig* (FB-IL) sahen, eine Bewertung vorzunehmen (Tabelle 70). Starke Diskrepanzen zwischen der *Einschätzung des eigenen Fachwissens* (NW-SF) und der selbst eingeschätzten *Fähigkeit zur Bewertung* (FB), wie in Studie S2.I1, konnten nicht festgestellt werden. Auch hier wurde die Selbsteinschätzung der Bewertung verhaltener, was daran erkenntlich wurde, dass sich die Teilnehmenden im Post-Test *nicht in der Lage* (FB-NL) sahen, eine Positionierung vorzunehmen. Bei zwei Teilnehmenden konnte zudem im Post-Interview keine Kodierung vorgenommen werden, da keine Aussagen mit hinreichenden inhaltlichen Bezugspunkten vorhanden waren.

Tabelle 70: S3.I1&2 - Ergebnisse der Interviewerhebung, *Fähigkeit zur Bewertung (FB)*, *N* = 10

Kat.	1. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
FB: Fähgkeit zur Bewertung	FB-NL: Nicht in der Lage	6	8	8	12	**N=0**
	FB-UE: Unter Einschränkungen in der Lage	2	0	2	0	**N=1**
	FB-IL: In der Lage	2	0	2	0	**N=2**
Summe FB-UE/IL:				**4**	**0**	
Einzelzählung FB-UE/IL:		**4**	**0**			

Eine direkte Gegenüberstellung der Ergebnisse kann der Tabelle 71 entnommen werden. Diese zeigt die Anzahl verschiedener Aspekte pro Proband*in pro *Teilbereich* (NW-A, EP-A und SR-A) auf. Im Bereich der Selbstreflexion wurden die beiden Subkategorien der *internen* (SR-A-IF) und *externen Faktoren* (SR-A-EF) nochmal separat dargestellt. Bei den Definitionen wurde die Anzahl verschiedener Teilbereiche der Grunddefinition gezählt, die eine informierte Definition widerspiegeln. Dies entsprach beim *Verständnis der Selbstreflexion* (SR-V) dem *expliziten Teilverständnis* (SR-V-ET) und beim Verständnis *Kritischen Denkens* (KD-V) den *Teilbereichen der Grunddefinition* (KD-V-TG).

Tabelle 71: S3.I1&2 - Ergebnisse der Interviewerhebung, Übersicht über die Einzelzählungen der verschiedenen Teilbereiche, $N = 10$

	Einzelzählung	NW-A		SR-A		EP-A		KD-V		SR-V		SR-A-IF		SR-A-EF	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Teilnehmende aus S2	XJ36JE	3	1	4	7	0	3	0	1	2	2	2	5	2	2
	BN10LB	3	5	5	11	1	5	0	3	2	3	3	8	2	3
	XF21FM	1	1	3	7	0	2	0	3	0	2	2	6	1	1
	XJ27BO	3	6	1	5	0	2	0	1	0	1	1	3	0	2
Teilnehmende aus S3	SJ16DI	3	5	5	5	1	4	0	3	0	2	4	5	1	0
	SV32BT	1	5	4	8	1	7	0	3	1	2	1	5	3	3
	IP17SE	3	6	3	6	0	3	1	3	1	1	2	5	1	1
	EF28SH	4	5	8	10	1	7	1	2	2	1	5	7	3	3
	WJ32FE	3	4	4	5	0	4	0	2	0	2	1	3	3	2
	XJ19MF	2	3	5	9	1	1	0	2	1	2	4	7	1	2
	MW	2,600	4,100	4,200	7,300	0,500	3,800	0,200	2,300	0,900	1,800	2,500	5,400	1,700	1,900
	Stabw.	0,966	1,853	1,814	2,163	0,527	2,044	0,422	0,823	0,876	0,632	1,434	1,647	1,059	0,994
	Varianz	0,933	3,433	3,289	4,678	0,278	4,178	0,178	0,678	0,767	0,400	2,056	2,711	1,122	0,989
	Max	4	6	8	11	1	7	1	3	2	3	5	8	3	3
	Min	1	1	1	5	0	1	0	1	0	1	1	3	0	0
	Median	3	5	4	7	0,5	3,5	0	2,5	1	2	2	5	1,5	2
	Modus	3	5	4	5	0	3	0	3	0	2	2	5	1	2

Auf Grundlage dieser Ergebnisse wurde mit jedem Teilbereich ein Signifikanztest durchgeführt, um signifikante Veränderungen zwischen Pre- und Post-Test zu identifizieren (Tabelle 72).

Tabelle 72: S3.I1&2 - Ergebnisse der Interviewerhebung, Wilcoxon-Test der Einzelzählungen der verschiedenen Teilbereiche, $N = 10$

S3.I1&2 - Wilcoxon-Test

Teststatistiken[a] Einfachzählung	NW_A_Post - NW_A_Pre	SR_A_Post - SR_A_Pre	EP_A_Post - EP_A_Pre	KD_V_Post - KD_V_Pre	SR_V_Post - SR_V_Pre	SR_A_IF_Post - SR_A_IF_Pre	SR_A_EF_Post - SR_A_EF_Pre
Z	-2,090b	-2,692b	-2,682b	-2,836b	-2,165b	-2,820b	-,707b
Asymp. Sig. (2-seitig)	0,037	0,007	0,007	0,005	0,030	0,005	0,480
Exakte Sig. (2-seitig)	0,039	0,004	0,004	0,002	0,047	0,002	0,750
Exakte Sig. (1-seitig)	0,020	0,002	0,002	0,001	0,023	0,001	0,375
Punkt-Wahrscheinlichkeit	0,006	0,002	0,002	0,001	0,020	0,001	0,219
a. Wilcoxon-Test							
b. Basiert auf negativen Rängen.							
Effektstärke	0,467	0,602	0,600	0,634	0,484	0,631	0,158

Alle Teilbereiche zeigten bei einem Signifikanzniveau $\alpha = 0{,}05$ statistisch signifikante Unterschiede. Während der Teilbereich des *naturwissenschaftlichen Wissens* (NW-A) mit $z = -2{,}090$, $p = 0{,}039$ und einer mittleren Effektstärke von $r = 0{,}467$ nahe an

einem starken Effekt lag, überschritten die anderen beiden Teilbereiche, die *Selbstreflexion* (SR-A) mit $z = -2{,}692$, $p = 0{,}004$ sowie das *epistemologische Verständnis* (EP-A) mit $z = -2{,}682$, $p = 0{,}004$, den Schwellenwert für einen starken Effekt ($r = 0{,}5$) deutlich. Hervorzuheben ist, dass die Perspektiverweiterung im Bereich der *Selbstreflexion* (SR-A) vorwiegend im Bereich der *internen Faktoren* (SR-A-IF) stattfand.

Auch das Verständnis der Begrifflichkeiten der *Selbstreflexion* (SR-V) und des *Kritischen Denkens* (KD-V) zeigte signifikante Unterschiede zwischen Pre- und Post-Test. Während das *Verständnis der Selbstreflexion* (SR-V) bei $z = -2{,}165$, $p = 0{,}047$ eine mittlere Effektstärke von $r = 0{,}484$ aufwies, zeigte die *Verständniserweiterung im Bereich des Kritischen Denkens* (KD-V) eine deutliche Signifikanz mit $z = -2{,}836$, $p = 0{,}047$ und einer Effektstärke von $r = 0{,}634$.

Diese Veränderungen lassen sich auch grafisch veranschaulichen, indem die absolute Anzahl der Aspekte aus den Einzelzählungen durch die maximale Anzahl der Kategorien des jeweiligen Leitfadens geteilt wird. Das Ergebnis ist ein relativer Wert, der die Vergleichbarkeit zwischen den Leitfäden gewährleistet und eine konsistente Darstellung in einem Netzdiagramm ermöglicht. Auf diese Weise wird ersichtlich, wie viel Prozent der maximal möglichen Aspekte in jedem Teilbereich erreicht wurden. Für eine summative Übersicht wurden die Mittelwerte aller Teilnehmenden zur Darstellung herangezogen (Abbildung 22). Schwerpunkt dieser Darstellung liegt auf den drei Teilbereichen der *Bewertungsaspekte (*NW-A, SR-A, EP-A) sowie auf dem Verständnis der Begrifflichkeiten *Selbstreflexion* (SR-V) und *Kritisches Denken* (KD-V). Diese fünf Bereiche wurden in den Fokus der Untersuchung gerückt, da sie das zentrale Untersuchungsinteresse repräsentieren und sich skalierbar darstellen lassen. Die Zusatzfragestellungen zur *fachlichen Bewertung* (NW-FB), *Selbsteinschätzung des Fachwissens* (NW-SF) sowie zur *Fähigkeit zur Bewertung* (FB) wurden bei dieser Darstellung nicht berücksichtigt

Die triadische Darstellung rekurriert auf das Triadische Modell Kritischen Denkens und bildet sowohl die Perspektiven der einzelnen Teilbereiche (SR-A, EP-A und NW-A) als auch die Zusammenführung der empirisch gewonnenen Daten mit den konzeptionellen Ansätzen ab. Die fünfteilige Darstellung erweitert dies um das Verständnis der in diesem Projekt relevanten Begrifflichkeiten *Kritisches Denken* (KD-V) und *Selbstreflexion* (SR-V).

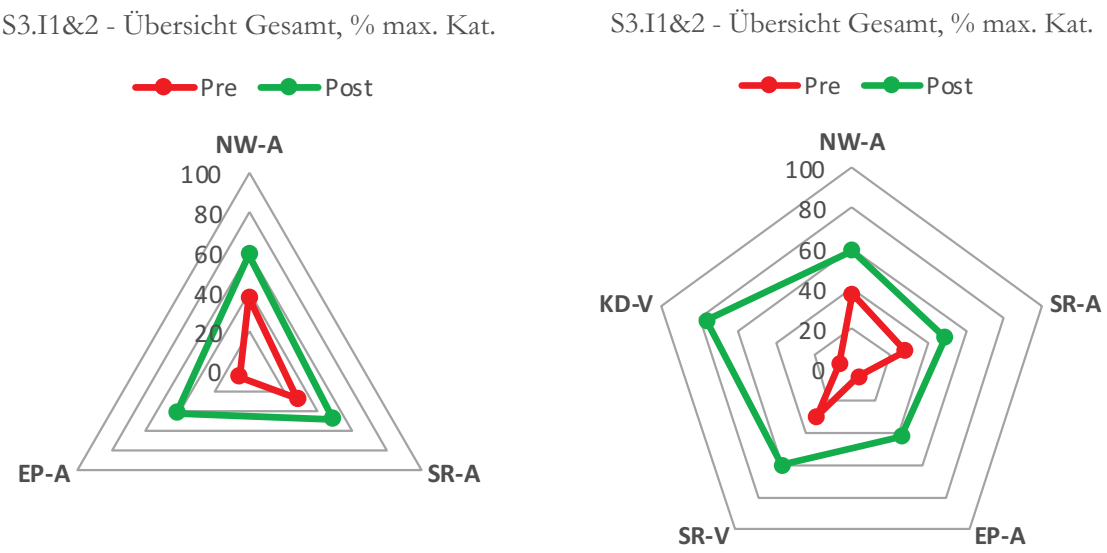
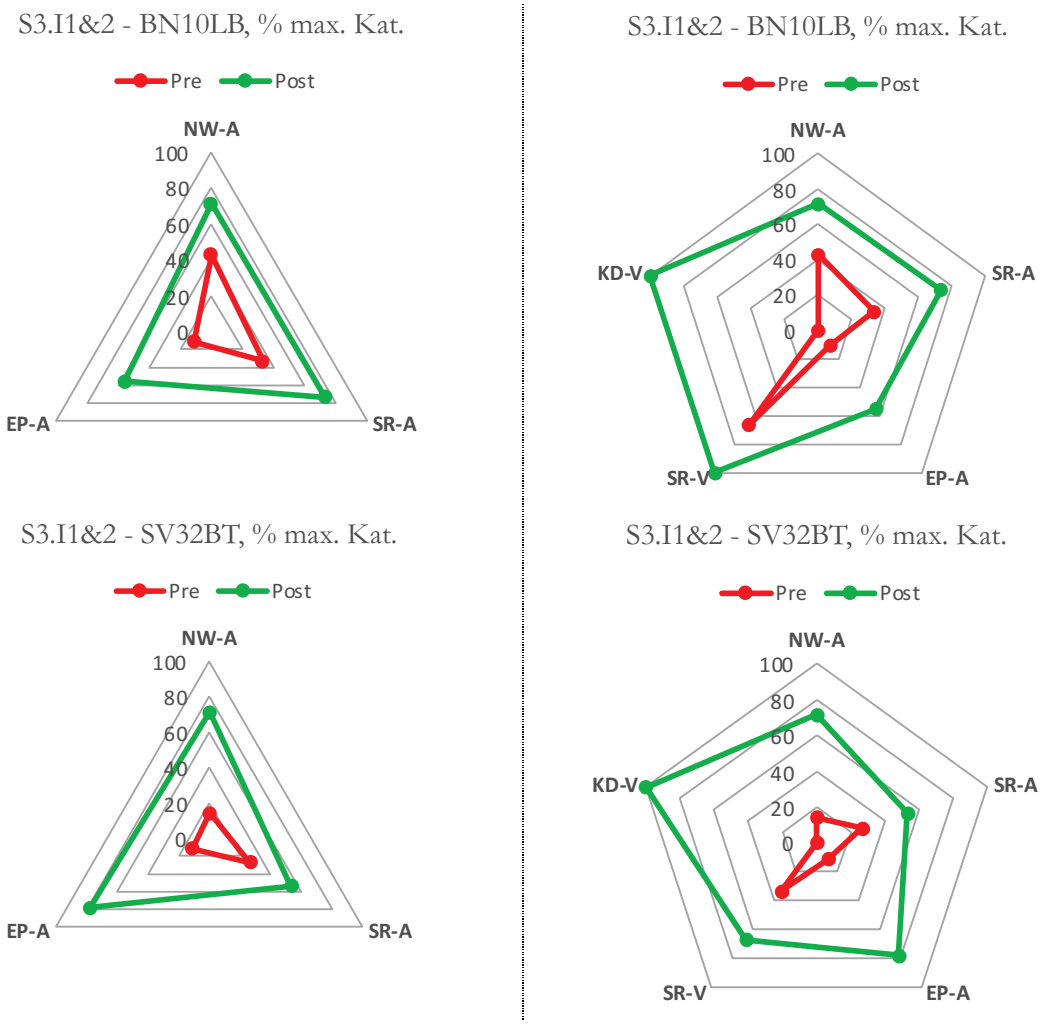


Abbildung 22: S3.I1&2 - Gesamtübersicht über die prozentualen Anteile an den einzelnen Teilbereichen (% max. Kat. bezeichnet den prozentualen Anteil an den maximal möglichen Kategorien)

Diese Darstellung kann auch individuell für jede*n Teilnehmende*n angewendet werden, um die Veränderungen zwischen Pre- und Post-Befragung auf Ebene einzelner Personen aufzuzeigen (Abbildung 23).



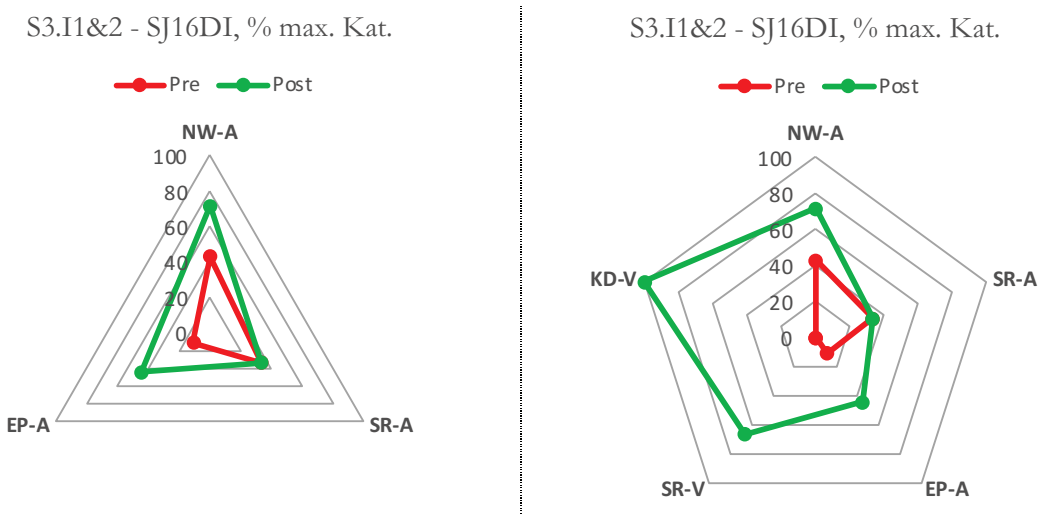


Abbildung 23: S3.I1&2 - Exemplarische Auswahl individueller Vergleichsdiagramme
(% max. Kat. bezeichnet den prozentualen Anteil an der maximalen Anzahl möglicher Kategorien)

Abbildung 23 zeigt drei exemplarische Vergleichsdiagramme, in denen die Veränderungen zwischen Pre- und Post-Interview durch die unterschiedlichen Ausprägungen im Netzdiagramm ersichtlich werden.

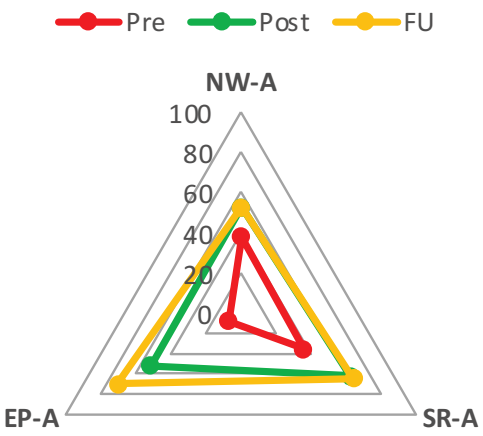
Neben den Pre- und Post-Interviews wurde mit drei freiwilligen Proband*innen etwa sechs Monate später ein Follow-Up-Interview durchgeführt, um die Nachhaltigkeit der gemachten Veränderungen zu überprüfen. Hierfür wurden zwei neue Kontexte (Kontext C: Plastik und Kontext D: Fracking) verwendet und diese den Teilnehmenden randomisiert zugeteilt. Weitere Informationen zu den Kontexten sind in den zugehörigen Kapiteln (4.2.6.4 und 4.2.6.5) zu finden.

Die Ergebnisse des Follow-Up-Interviews waren weitgehend mit den Ergebnissen der Post-Interviews vergleichbar (Abbildung 24). Die Anzahl verschiedener Aspekte im Bereich *naturwissenschaftlicher Bewertung* (NW-A) blieb konstant hoch und auch die Ergebnisse in den Bereichen der *Selbstreflexion* (SR-A) und des *epistemologischen Verständnisses* (EP-A) zeigten ähnlich hohe Werte wie im Post-Interview. Zwar traten in vereinzelten Fällen Reduktionen bezüglich der Anzahl hinzugezogener Perspektiven auf, diese blieben jedoch auf eine Differenz von 1 begrenzt. Im Gegenteil wurden in den Follow-Up-Interviews häufiger höhere Werte und differenziertere Perspektiven festgestellt als im Post-Interview. Auch das Verständnis der Begriffe *Selbstreflexion* (SR-V) und *Kritisches Denken* (KD-V) blieb in allen Fällen auf der höchsten Niveaustufe (Niveau 2). Dies deutet weiterhin auf ein fundiertes und informiertes Verständnis dieser Begriffe hin.

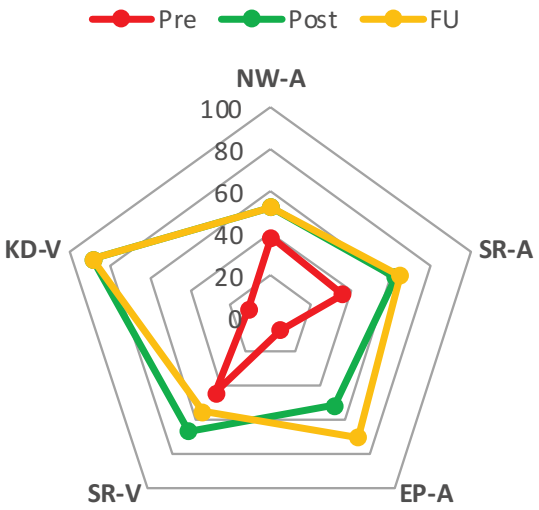
Aufgrund des positiven und strukturierten Verlaufs der Interviews wurden nach dem Interview gezielte Nachfragen gestellt, um die Hintergründe für die Reflexionen der

Teilnehmenden zu ermitteln. Dabei gaben die Teilnehmenden an, dass die Auseinandersetzung mit der Thematik des Kritischen Denkens sie sensibler gemacht habe, insbesondere in Bezug auf die Wahrnehmung von Informationen und Bewertungsprozessen. Diese gesteigerte Sensibilität begleite sie weiterhin im Alltag, insbesondere in Entscheidungssituationen.

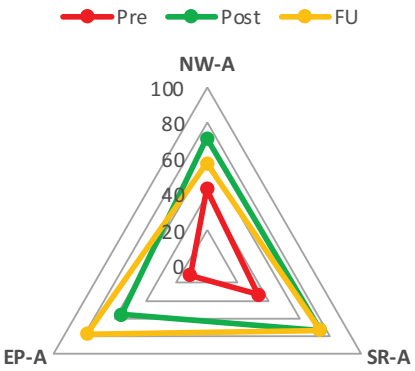
S3.I1&2 - Übersicht Follow Up, % max. Kat.



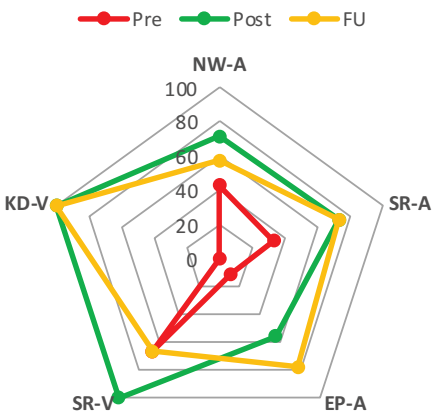
S3.I1&2 - Übersicht Follow Up, % max. Kat.



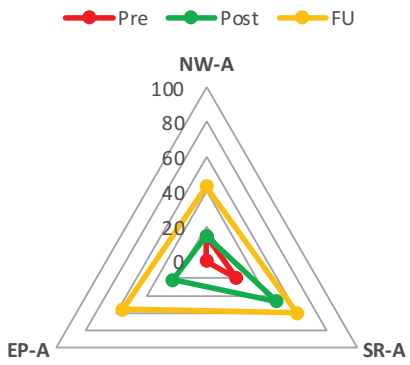
S3.I1&2 - BN10LB, % max. Kat.



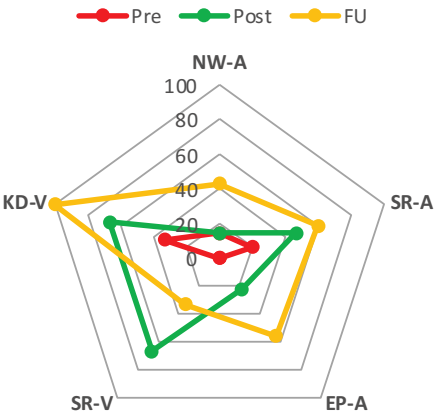
S3.I1&2 - BN10LB, % max. Kat.



S3.I1&2 - XF21FM, % max. Kat.



S3.I1&2 - XF21FM, % max. Kat.



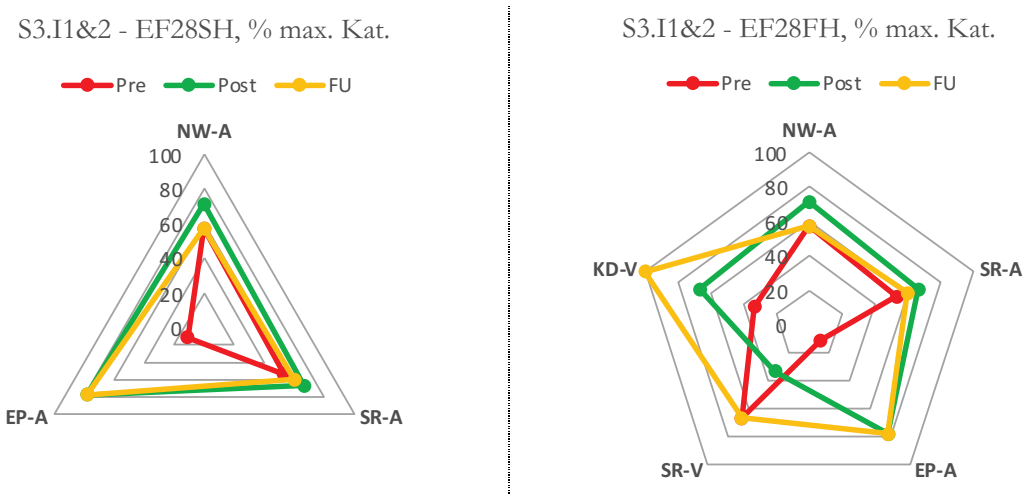


Abbildung 24: S3.I1&2 + FU - Gesamtübersicht und Vergleichsdiagramme der Teilnehmenden: Prozentuale Anteile an den einzelnen Teilbereichen (% max. Kat. bezeichnet den prozentualen Anteil an der maximalen Anzahl möglicher Kategorien)

Diese Ergebnisse in Verbindung mit der Rückmeldung der Teilnehmenden deuten auf eine Nachhaltigkeit der erzielten Resultate der Vermittlungseinheit hin und liefern erste Anzeichen für die Funktionalität des Instruments. Dies legt nahe, dass das Instrument nicht nur dazu geeignet ist, Kritisches Denken in der durch die Semineareinheiten vermittelten Form abzubilden, sondern weist auch darauf hin, dass Kritisches Denken als multiperspektivischer Denkansatz gezielt gefördert werden kann.

Aufgrund der geringen Stichprobenzahl von $N = 3$ für die Follow-Up-Stichprobe wurde auf eine statistische Auswertung verzichtet.

5.3.2 Fragebogenerhebung S3.F1&2

Der in Studie S3.F1&2 eingesetzte Fragebogen stellte eine Erweiterung der vorangegangenen Fragebogenformate aus Studie S1 und S2 dar. Zusätzlich zu den bereits etablierten Items der SRIS von Silvia (2021) und des CRT von Frederick (2005) wurden erstmals zusätzliche Items integriert, die das Verständnis der Begriffe der Selbstreflexion und des Kritischen Denkens sowie die beeinflussenden Faktoren in Bewertungsprozessen untersuchen sollten. Ziel war es, tiefere Einblicke in die Bewertungsprozesse und Begriffsverständnisse zu gewinnen sowie die Praktikabilität eines erweiterten Fragebogenformats für den Einsatz in größeren Stichproben zu erproben.

Der erste Teil des Fragebogens bestand aus den obligatorischen Fragen zur befragten Person und leitete die Genese des Codes an. Darauf folgte der erste neue Block der *Verständnisfragen* (VF), welcher im offenen Fragenformat das Verständnis der Begrifflichkeiten und Einflussfaktoren erfragte (Tabelle 73).

Tabelle 73: S3.F1&2 - Fragestellungen des Blocks *Verständnisfragen* (VF)

Fragen-code	Fragestellung	Itemformat
VF1	Was verstehen Sie unter dem Begriff *Selbstreflexion*?	Offen
VF2	Was verstehen Sie unter *Kritischem Denken*?	Offen
VF3	Was glauben Sie, was beeinflusst Sie besonders, wenn Sie im Alltag Informationen bewerten, die einen naturwissenschaftlichen Bezug haben?	Offen

Auf den Block der *Verständnisfragen* (VF) folgte ein Abschnitt, dessen Fokus auf der Einschätzung der *Introspektion* (IS) der Teilnehmenden lag. Ziel war die Ergründung, inwieweit Chemielehramtsstudierende bewusst interne Faktoren bei der Bewertung naturwissenschaftlicher Kontexte reflektieren. Die Befragten konnten ihr selbstreflexiven Verhalten auf einer unipolaren fünfstufigen Likert-Skala einstufen, wobei die Skala von 1 = *nie/gar nicht* bis 5 = i*mmer/sehr ausführlich* reichte (Tabelle 74). Methodische Details zur Itemkonstruktion sind im Kapitel 4.2.3 einzusehen.

Tabelle 74: S3.F1&2 - Fragestellungen des Blocks *Introspektion* (IS)

Fragen-code	Fragestellung	Itemformat
IS1	Wie oft machen Sie sich Ihre eigenen Gedanken und Gefühle bewusst, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	Geschlossen
IS2	Inwieweit setzen Sie sich mit Ihren eigenen Moral-und Wertvorstellungen auseinander, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	Geschlossen
IS3	Wie häufig denken Sie über Ihre persönlichen Erfahrungen und Hintergründe nach, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	Geschlossen
IS4	Inwiefern reflektieren Sie Ihre Interessen und Motivationen, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	Geschlossen
IS5	Wie häufig hinterfragen Sie Ihre eigenen Vorurteile und Annahmen, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	Geschlossen

Die fünf Items wurden so gestaltet, dass sie unterschiedliche Facetten der *Introspektion* (IS) abbilden. Diese umfassten die Reflexion eigener Gedanken und Gefühle, moralischer Werte, persönlicher Erfahrungen, Interessen sowie bestehender Vorurteile.

Der anschließende Block *Einflussfaktoren* (EF) diente dazu, sowohl die von den Teilnehmenden wahrgenommenen und reflektierten Einflussfaktoren mithilfe geschlossener als auch offener Frageformate zu erheben. Die Fragestellungen wurden in einer Abfolge vom Allgemeinen zum Expliziten formuliert, um zunächst ein unbeeinflusstes Bild zu gewinnen, bevor spezifische Faktoren gezielt erhoben wurden (Tabelle 75).

Tabelle 75: S3.F1&2 - Fragestellungen des Blocks *Einflussfaktoren* (EF)

Fragencode	Fragestellung	Itemformat
EF1	Welche Faktoren haben Ihrer Meinung nach einen Einfluss auf Ihre Entscheidungen, wenn Sie ein Thema mit naturwissenschaftlichem Bezug bewerten sollen?	Offen
EF2	Inwieweit glauben Sie, dass Sie in der Lage sind, vornehmlich auf fachlicher Ebene zu urteilen?	Geschlossen
EF3	Was wissen Sie über kognitive Verzerrungen?	Offen
EF4	Wie sehr hinterfragen Sie folgende Faktoren, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	Geschlossen
	Personenbezogene Faktoren: EF4_1 Eigene Persönlichkeit; EF4_2 Stimmung und Emotion; EF4_3 Werte und Moralvorstellungen; EF4_4 Eigene Grundeinstellung; EF4_5 Eigene Bildung; EF4_6 Eigenes Alter; EF4_7 Eigenes Geschlecht; EF4_8 Eigene soziale Schicht; EF4_9 Eigene Religion; EF4_10 Eigene Ethnie; EF4_11 Eigene Kultur; EF4_12 Wissen; EF4_13 Medialer Kontext; EF4_14 Politische Einstellung; EF4_15 Erfahrungen und Erinnerungen; EF4_16 Selbstkonzept und Selbstwertgefühl	
	Bezug zu kognitiven Verzerrungen: EF4_17 Soziale Normen und Erwartungen; EF4_18 Motivation und Bedürfnisse; EF4_19 Verfügbarkeit von Alternativen; EF4_20 Intuition und Bauchgefühl; EF4_21 Gruppendynamik und Konformitätsdruck; EF4_22 Erwartungen und Vorurteile; EF4_23 Framing und Präsentation von Information; EF4_24 Selektive Wahrnehmung und Bestätigungsfehler; EF4_25 Übergeneralisierung und Kategorisierung; EF4_26 Komplexität der Information	

Die Items EF2 und EF4 wurden auf fünfstufigen Likert-Skalen bewertet, die an die jeweilige Fragestellung angepasst wurden. Während EF2 von 1 = *überhaupt nicht* bis 5 = *vollständig* skaliert war, reichten die Likert-Skalen der Items EF4 von 1 = *gar nicht* bis 5 = u*mfassend*. Die Items von EF4 umfassten insgesamt 26 Einflussfaktoren, welche in 16 personenbezogene Faktoren und zehn Faktoren mit Bezug zu kognitiven Verzerrungen unterteilt werden konnten.

Obwohl die Frage EF3 in ihrer Formulierung den *Verständnisfragen* (VF) ähnelt, wurde der Bereich der kognitiven Verzerrungen thematisch zu den *Einflussfaktoren* (EF) gezählt und diente als offene Fragestellung zudem der Auflockerung des ansonsten geschlossenen Fragenformats dieses Blocks.

Der letzte Fragenblock *Bewertung und Gewichtung* (BG) zielte darauf ab, die Einschätzung der empfundenen Relevanz der in EF4 genannten Faktoren sowie den Grad, in dem die Teilnehmenden diese Faktoren bewusst in ihre Bewertungen einbeziehen, zu erfassen (Tabelle 76). Die Likert-Skala dieser geschlossenen Items reichte von 1 = *nie/gar nicht* bis zu 5 = *immer/sehr stark*, wodurch die subjektive Einschätzung der Teilnehmenden hinsichtlich des Einflusses von Faktoren, wie Emotionen oder Werte, im Vergleich zum Fachwissen erfasst wurde.

Tabelle 76: S3.F1&2 - Fragestellungen des Blocks *Bewertung und Gewichtung* (BG)

Fragen-code	Fragestellung	Itemformat
BG1	Wie stark schätzen Sie den Einfluss der vorhin genannten Faktoren (Bauchgefühl, Stimmung usw.) im Vergleich zum Fachwissen ein, wenn Sie ein Thema mit naturwissenschaftlichem Bezug bewerten sollen?	Geschlossen
BG2	Inwieweit wägen Sie diese Faktoren (Bauchgefühl. Stimmung usw.) bewusst gegen Ihr Fachwissen ab, wenn Sie ein Thema mit naturwissenschaftlichem Bezug bewerten sollen?	Geschlossen

Die Items der bereits etablierten Testinstrumente des CRT und SRIS wurden zwischen die einzelnen Fragenblöcke gesetzt und die Reihenfolge der CRT-Fragen randomisiert. Dies diente dazu, Ablenkungen, Konsistenz- und Ankereffekte zu reduzieren. Der finale Aufbau des eingesetzten Fragebogens kann Tabelle 77 entnommen werden.

Der finale Fragebogen der Studie S3.F1&2 umfasste somit fünf offenen Fragestellungen, 46 geschlossene Items, die mittels einer Likert-Skala evaluiert wurden, sowie die drei Aufgaben des Cognitive Reflection Tests (CRT) nach Frederick (2005). Die Länge des Fragebogens wurde bewusst gewählt, um sowohl eine hinreichende Tiefe und Breite der Erhebung sicherzustellen als auch die Belastung der Teilnehmenden

in einem akzeptablen Rahmen zu halten. Der überwiegende Anteil geschlossener Likert-Items trug dazu bei, die Bearbeitungszeit effizient zu gestalten, sodass diese mit etwa 20 bis 30 Minuten in einem akzeptablen Rahmen blieb. Überdies wurden durch eine klar strukturierte Anordnung der Fragen sowie eine thematische Gliederung Maßnahmen ergriffen, um die kognitive Belastung zu minimieren und die Datenqualität sicherzustellen.

Tabelle 77: S3.F1&2 - Abfolge der Frageblöcke im finalen Fragebogen

Fragencode	Fragestellung	Itemformat
PI	Persönliche Informationen	
VF1	Verständnisfrage 1	Offen
VF2	Verständnisfrage 2	Offen
CRT1	Erste randomisierte Frage des CRT	Zahleneingabe
VF3	Verständnisfrage 3	Offen
IS	Introspektion	Geschlossen
SR	SRIS – Self-Reflection	Geschlossen
CRT2	Zweite randomisierte Frage des CRT	Zahleneingabe
IN	SRIS - Insight	Geschlossen
EF1	Einflussfaktoren 1	Offen
CRT3	Dritte randomisierte Frage des CRT	Zahleneingabe
EF2-4	Einflussfaktoren 2-4	Offen/Geschlossen
BG	Bewertung und Gewichtung	Geschlossen

Die Gestaltung des Umfangs des Fragebogens basierte auf der Zielsetzung, eine umfassende und differenzierte Datengrundlage für die Untersuchung zu schaffen. Während die offenen Fragen dazu dienten, qualitative Einblicke in die Denk- und Reflexionsprozesse der Teilnehmenden zu gewinnen, ermöglichten die geschlossenen Likert-Items eine standardisierte Erhebung quantitativer Daten, die eine vergleichende Analyse und statistische Auswertung begünstigten. Der CRT ergänzte das Instrumentarium, indem er die kognitiven Reflexionsfähigkeiten der Teilnehmenden erfasste und somit eine zusätzliche Dimension für die Interpretation der Ergebnisse bereitstellte. Die leitenden Untersuchungsfragen der einzelnen Fragenblöcke sind in Tabelle 78 aufgeführt.

Der eingesetzte Fragebogen wurde vor seinem Ersteinsatz mit $N = 4$ Masterstudierenden pilotiert sowie zur Optimierung zusätzlich mit einem Doktoranden der Chemiedidaktik durchgeführt. Die Beantwortung nahm etwa 20 bis 30 Minuten in Anspruch, was die Praktikabilität des Instruments im Hinblick auf dessen Länge und Struktur bestätigte.

Tabelle 78: S3.F1&2 – Leitende Untersuchungsfragen der einzelnen Fragenblöcke

Fragenblock		Zugehörige Untersuchungsfrage
Offene Fragestellungen	VF1: *Verständnis Selbstreflexion* (VS)	Welches Verständnis haben Chemielehramtsstudierende von dem Begriff der Selbstreflexion? (VS)
	VF2: *Verständnis Kritisches Denken* (VK)	Welches Verständnis haben Chemielehramtsstudierende von dem Begriff des Kritischen Denkens? (VK)
	VF3 & EF1: *Bewertungsfaktoren* (BF)	Welche Faktoren werden von Chemielehramtsstudierenden als beeinflussend bei der Bewertung von Themen mit naturwissenschaftlichem Bezug wahrgenommen?
	EF3: *Verständnis kognitive Verzerrungen* (KV)	Welches Verständnis haben Chemielehramtsstudierende von dem Begriff der kognitiven Verzerrungen?
Geschlossene Fragestellungen	IS: *Introspektion*	Inwieweit reflektieren Chemielehramtsstudierende bewusst interne Faktoren bei der Bewertung naturwissenschaftlicher Kontexte?
	EF2: *Einflussfaktoren*	Inwieweit sehen sich Chemielehramtsstudierende in der Lage, vornehmlich auf fachlicher Ebene zu urteilen?
	EF4: *Einflussfaktoren*	Welche Faktoren werden von Chemielehramtsstudierenden bei der Bewertung von naturwissenschaftlichen Themen besonders reflektiert?
	BG: *Bewertung und*	Inwieweit berücksichtigen Chemielehramtsstudierende interne und externe Einflussfaktoren bei der Bewertung naturwissenschaftlicher Kontexte, zusätzlich zum Fachwissen?
	SRIS: *Self-reflection and insight scale*	Inwieweit zeigen Chemielehramtsstudierende die Tendenz, selbstreflexives Verhalten zu zeigen?
	CRT: *Cognitive Reflection Test*	Inwieweit sind Chemielehramtsstudierende in der Lage, ihren intuitiven Erstimpuls bei Fragestellungen zu relativieren und eine rationale Problembetrachtung vorzunehmen?

5.3.2.1 Auswertung

Die Antworten auf die offenen Fragestellungen der Fragebogenerhebung S3.F1&2 wurden ebenfalls auf Grundlage der parallel für die Interviews entwickelten Kategorienleitfäden mittels der qualitativen Inhaltsanalyse nach Kuckartz (2016) ausgewertet. Die Kodierleitfäden der Fragebögen und Interviews basierten auf denselben theoretischen Grundlagen und wurden so gestaltet, dass sie inhaltlich konsistent sind.

Es wurden lediglich punktuelle Anpassungen vorgenommen, um den Anforderungen der verschiedenen Erhebungsinstrumente gerecht zu werden.

Aus diesem Grund werden an dieser Stelle lediglich Abweichungen von den bereits vorgestellten Kategoriensystemen thematisiert, um Redundanzen zu vermeiden. Zur besseren Unterscheidung zwischen den in den Interviews und den Fragebögen verwendeten Kategorien wurden die Kürzel, auch bei Kodierleitfäden mit hoher Übereinstimmung, an das jeweilige Instrument angepasst. Die vollständigen Kodierleitfäden für die offenen Fragestellungen sind zur Einsicht dem Anhang (A.3.1.2) beigefügt.

Zur Auswertung der *Verständnisfrage* 1 (VF1), welche das Verständnis des *Begriffs der Selbstreflexion* (VS) erhob, blieb das Kategoriensystem identisch zum Kodierleitfaden SR-V der Interviewerhebung. Lediglich die Kategorie der *Introspektion* (VS-ET-IS) wurde aufgrund der größeren Stichprobe und der Vielfalt der Antworten weiter ausdifferenziert, um die Zuordnung zu dieser Kategorie nachvollziehbarer und transparenter zu gestalten (Tabelle 79).

Tabelle 79: S3.F1&2 - Erweiterte Kategorien der *Introspektion* (VS-ET-IS) zur Auswertung von Frage VF1

Überkategorie	Kategorie	Subkategorie	Definition
VS-ET: *Explizites Teilverständnis*	VS-ET-IS: *Introspektion*	VS-ET-IS-AL: *Introspektion Allgemein*	Selbstreflexion wird als die Fähigkeit zur Introspektion verstanden.
		VS-ET-IS-PK: *Persönlichkeit*	Selbstreflexion wird als die Auseinandersetzung mit der eigenen Persönlichkeit verstanden.
		VS-ET-IS-GE: *Gedanken*	Selbstreflexion wird als die Auseinandersetzung mit den eigenen Gedanken verstanden.
		VS-ET-IS-EM: *Emotionen*	Selbstreflexion wird als die Auseinandersetzung mit den eigenen Emotionen verstanden.
		VS-ET-IS-WI: *Wissen*	Selbstreflexion wird als die Auseinandersetzung mit oder Bewertung von eigenem Wissen verstanden.
		VS-ET-IS-AN: *Ansichten/ Meinung*	Selbstreflexion wird als die Auseinandersetzung mit den eigenen Ansichten oder der eigenen Meinung verstanden.
		VS-ET-IS-MW: *Moral/Werte*	Selbstreflexion wird als die Auseinandersetzung mit den eigenen Moral- oder Wertvorstellungen verstanden.
		SR-V-ET-FA: *Fähigkeiten*	Selbstreflexion wird als die Auseinandersetzung mit den oder die Bewertung der eigenen Fähigkeiten verstanden.
		VS-ET-IS-EF: *Erfahrungen*	Selbstreflexion wird als die Auseinandersetzung mit den eigenen Erfahrungen verstanden.

Auch der Kodierleitfaden zur Auswertung der Verständnisfrage 2 (VF2), welche analog zur VF1 das Verständnis des Begriffs des Kritischen Denkens adressierte, blieb bis auf eine Anpassung deckungsgleich mit dem Kodierleitfaden KD-V aus der Interviewerhebung. Lediglich in der Kategorie *Teilbereich der Grunddefinition* (VK-TG) wurde die Subkategorie der *Multiperspektivität* (MP) ergänzt, da sich Antworten in den Fragebögen darauf bezogen und im Vergleich zu den Interviews keine Möglichkeit bestand, durch Nachfragen eine Ausdifferenzierung der Multiperspektivität zu erreichen.

Die Auswertung der Antworten auf die Fragen VF3 und EF1, welche beide die in Bewertungssituationen wahrgenommenen Einflussfaktoren feststellen sollten, wurde auf Grundlage der in der Auswertung der Interviewerhebung S3.I1&2 entwickelten Kodierleitfäden für die *Aspekte des naturwissenschaftlichen Wissens* (NW-A), die *Aspekte der Selbstreflexion* (SR-A) und die *Aspekte epistemologischen Verständnisses* (EP-A) durchgeführt und für das Fragebogenformat in einem einzelnen Kodierleitfaden für *Bewertungsfaktoren* (BF) zusammengefasst. Darüber hinaus verlief die Auswertung analog zur Interviewerhebung S3.I1&2.

Die letzte offene Fragestellung EF3 erfragte das *Verständnis von kognitiven Verzerrungen* (KV). Um die Angemessenheit der Antworten einordnen zu können, wurde zunächst eine Definition für kognitive Verzerrungen geschaffen, von der sich dann die Kategorien deduktiv ableiten ließen:

> Kognitive Verzerrungen sind systematische Denk- und Wahrnehmungsfehler, die dazu führen, dass Informationen fehlerhaft interpretiert, verarbeitet oder erinnert werden. Sie können sowohl durch interne Faktoren, wie individuelle Erfahrungen, Überzeugungen oder emotionale Zustände, als auch durch externe Einflüsse, wie soziale Normen oder die Präsentation von Informationen, bedingt sein. Diese Verzerrungen treten häufig in unsicheren oder komplexen Entscheidungskontexten auf und beeinträchtigen die Fähigkeit, objektive Urteile zu fällen. Sie haben weitreichende Auswirkungen auf das menschliche Denken und Verhalten, da sie Fehlurteile, falsche Erinnerungen oder verzerrte Wahrnehmungen begünstigen (basierend u. a. auf Gigerenzer, 1991; Gilovich et al., 2002; Kahneman, 1973; R. Pohl, 2004).

Vergleichbar zu den anderen Definitionsfragen wurde dadurch eine Abstufung in zwei Niveaus möglich (N = 0 und N = 1) (Tabelle 80).

Tabelle 80: S3.F1&2 – Kategorien zum *Verständnis von kognitiven Verzerrungen* (KV) zur Auswertung von Frage EF3

Überkategorie	Kategorie	Subkategorie	Definition	Niveau
KV: *Verständnis kognitive Verzerrungen*	KV-NA: *Naive Vorstellungen*	KV-NA-UK: *Unkonkret*	Selbstreflexion wird als die Fähigkeit zur Introspektion verstanden.	N = 0
		KV-NA-FV: *Fehlvorstellungen*	Selbstreflexion wird als die Auseinandersetzung mit der eigenen Persönlichkeit verstanden.	
	KV-TG: *Teilbereich der Grunddefinition*	KV-TG-DF: *Denk- und Wahrnehmungsfehler*	Selbstreflexion wird als die Auseinandersetzung mit den eigenen Emotionen verstanden.	N = 1
		KV-TG-IF: *Einfluss von internen Faktoren*	Selbstreflexion wird als die Auseinandersetzung mit oder Bewertung des eigenen Wissens verstanden.	
		KV-TG-EF: *Einfluss von externen Faktoren*	Selbstreflexion wird als die Auseinandersetzung mit den eigenen Ansichten oder der eigenen Meinung verstanden.	

Die Ergebnisse der Kodierungen der offenen Fragen wurden in Excel tabellarisch aufgetragen, um die Werte aus Pre- und Post-Erhebungen miteinander zu vergleichen. Neben einer qualitativen Auswertung wurde ebenfalls zur Analyse die Maße der zentralen Tendenz und die Streumaße berechnet und die Unterschiede zwischen beiden Erhebungszeitpunkten mittels eines Signifikanztests überprüft. Die eingesetzten Kodierleitfäden für die offenen Fragestellungen sind im Anhang (A.3.1.2) einzusehen.

Die Ergebnisse der geschlossenen Items des CRT, der SRIS sowie der Blöcke IS, EF2, EF4 und BG wurden in reine Zahlenformate überführt und invers formulierte Items, wie im SRIS-IN, rekodiert. Anschließend wurden auch hier die Maße der zentralen Tendenz und die Streumaße berechnet, um die Verteilung der Daten zu analysieren. Auf die Berechnung von Cronbach's Alpha für die einzelnen Itemblöcke wurde verzichtet, da die Items thematisch heterogene Konstrukte abbilden. Zur Überprüfung von Veränderungen und Unterschieden zwischen Gruppen wurde der Wilcoxon-Test als nichtparametrischer Signifikanztest eingesetzt.

Weiterführende Informationen zu den durchgeführten Rechenoperationen sind im zugehörigen Kapitel 4.3.2 einzusehen.

5.3.2.2 Ergebnisse

Die Antworten des Pre-Tests auf die Verständnisfrage 1 (VF) zum *Begriff der Selbstreflexion* (VS) zeigten bei 69,8 % (11 von 16) der Befragten ein *explizites Teilverständnis* (VS-ET), das sich primär auf das *Überdenken von Handlungen* (VS-ET-HA) konzentrierte (Tabelle 81). Zehn der 16 Befragten (62,5 %) fassten zudem auch Aspekte der *Introspektion* (VS-ET-IS) unter den Selbstreflexionsbegriff, wobei pro Proband*in immer nur eine Ausprägung genannt wurde. Vier Teilnehmende wiesen naivere Vorstellungen *mit implizitem Teilverständnis* (VS-IT) auf, die allesamt auf den Bereich der *Selbstverbesserung* (VS-NA-SV) entfielen.

Tabelle 81: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis Selbstreflexion* (VS), *N* = 16

Kat.	1.UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
VS: Verständnis Selbstreflexion	**VS-ND: Nicht definierbar (ND)**			0	0	0	0	**N=0**
	VS-IT: Implizites Teilverständnis		VS-IT-AB: Abwägen	0	0	0	0	**N=1**
			VS-IT-SV: Selbstverbesserung	4	0	4	0	
			VS-IT-EP: Empathie	0	0	0	0	
			VS-IT-HF: Hinterfragen	0	0	0	0	
	Summe VS-IT:					**4**	**0**	
	Einzelzählung VS-IT:			**4**	**0**			
	VS-ET: Explizites Teilverständnis	**VS-ET-IS: Introspektion**	VS-ET-IS-AL: Introspektion allg.	1	2	1	2	**N=2**
			VS-ET-IS-PK: Persönlichkeit	0	1	0	1	
			VS-ET-IS-GE: Gedanken	2	4	2	4	
			VS-ET-IS-EM: Emotionen	0	5	0	5	
			VS-ET-IS-WI: Wissen	3	3	3	3	
			VS-ET-IS-AN: Ansichten/Meinung	0	7	0	7	
			VS-ET-IS-MW: Moral/Werte	0	3	0	3	
			VS-ET-IS-FA: Fähigkeiten	3	0	3	0	
			VS-ET-IS-EF: Erfahrungen	1	0	1	0	
			VS-ET-HA: Handlungen überdenken	11	6	11	6	
			VS-ET-EX: Externe Prägungen	0	5	0	5	
	Summe					**21**	**36**	
	Einzelzählung			**6**	**9**			

Gegenüber dem Pre-Test zeigte sich im Post-Test eine Perspektiverweiterung hinsichtlich der Ausdifferenzierung des Verständnisses. Während im Pre-Test das *explizite Teilverständnis* (VS-ET) durchschnittlich 1,3 Teilbereiche umfasste, erweiterte sich das Verständnis der Teilnehmenden im Post-Test auf durchschnittlich 2,3 erfasste Teilbereiche. Das *Überdenken von Handlungen* (VS-ET-HA) wurde im Post-Test seltener als Bestandteil der Selbstreflexion genannt, während die Reflexion *externer Prägungen* (VS-ET-EX) von fünf Proband*innen erstmals als Teilbereich der Selbstreflexion berücksichtigt wurde. Alle 16 Teilnehmenden erreichten Niveaustufe 2 und naivere Ansichten, die nicht ein explizites Teilverständnis erkennen ließen, waren nicht mehr vorhanden.

Bezüglich des *Verständnisses von Kritischem Denken* (VK) zeigten die Antworten auf Verständnisfrage 2 (VF2) deutlich uninformierte Vorstellungen (Tabelle 82). 12 der 16 Teilnehmenden (75 %) zeigten *naive Definitionsansätze* (VK-NA), wobei 25 % (4 von 16) auf ein unkonkretes *Hinterfragen* (VK-NA-HF) und 31,3 % (5 von 16) auf eine grundlegend *skeptische Haltung* (VK-NA-SK) entfielen. 25 % der Befragten (4 von 16) zeigten zudem keine definierbare Vorstellung der Begrifflichkeit. Nur 6 der 16 teilnehmenden Personen (37,5 %) erreichten Niveaustufe 1 und schnitten *Teilbereiche der Grunddefinition* (VK-TG) an, wobei der Schwerpunkt auf dem kritischen Umgang mit *Wissen* (VK-TG-WI) lag. Keiner der Studierenden erreichte Niveaustufe 2, also das *Zusammenspiel mehrerer Teilbereiche der Grunddefinition* (VK-MT).

Im Post-Test zeigten sich hingegen deutlich informiertere Ansichten und 75 % der Befragten (12 von 16) erreichten in ihrem Verständnis Niveaustufe 1. Von den zehn Teilnehmenden, welche im Pre-Test auf Niveaustufe 0 verblieben waren, erreichten sechs in Post-Test mindestens Niveaustufe 1. Auf dieser ordneten sieben Proband*innen den bewussten Umgang mit *Wissen* (VK-TG-WI) dem Kritischen Denken zu, weitere sieben betrachteten die *Selbstreflexion* (VK-TG-SR) als dessen Teilbereich und sechs Personen bezogen auch das *epistemologische Verständnis* (VK-TG-EP) mit ein. Sieben Teilnehmende erreichten im Post-Test zudem die Niveaustufe 2, das *Zusammenspiel mehrerer Teilbereiche* (VK-MT).

Tabelle 82: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis Kritisches Denken* (VK), *N* = 16

Kat.	1.UK	2. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf.Pre	Summe Mehrf. Post	Niveau
VK: Verständis Kritisches Denken	VK-NA: Naive Vorstellung	VK-NA-ND: Nicht definierbar	4	1	4	1	N=0
		VK-NA-HF: Hinterfragen	4	3	4	3	
		VK-NA-SK: Skepsis	5	2	5	2	
		VK-NA-SV: Selbstkritik/-verbesserung	0	0	0	0	
		VK-NA-TK: Tadeln/Kritisieren	1	0	1	0	
		VK-NA-SD: Selbstständiges Denken	1	1	1	1	
	Summe VK-NA:				**15**	**7**	
	Einzelzählung VK-NA:		**15**	**7**			
	VK-TG: Teilbereich der Grunddefinition	VK-TG-MP: Multiperspektivität	0	2	0	2	N=1
		VK-TG-SR: Selbstreflexion	1	7	1	8	
		VK-TG-WI: Wissen	5	7	7	7	
		VK-TG-EP: Epistemologisches Verständnis	0	6	0	6	
	Summe VK-TG:				**8**	**23**	
	Einzelzählung VK-TG:		**6**	**22**			
	VK-MT: Zusammenspiel mehrerer Teilbereiche		**0**	**7**			N=2

Das im Fragenblock EF eingeholte *Verständnis kognitiver Verzerrungen* (KV) zeigte vergleichbar zum Kritischen Denken im Pre-Test weitestgehend *naive* (KV-NA) oder nicht vorhandene Begriffsdefinitionen (Tabelle 83). Nur vier Teilnehmende wiesen eine Vorstellung des Begriffs auf, welcher entweder als *Denk- und Wahrnehmungsfehler* (KV-TG-DF) oder *Einfluss interner Faktoren* (KV-TG-IF) charakterisiert wurde. Elf Personen konnten den Begriff demgegenüber *nicht definieren* (KV-NA-UK).

Tabelle 83: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis kognitive Verzerrungen* (KV), *N* = 16

Kat.	1.UK	2. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
KV: Verständnis kognitive Verzerrungen	KV-NA: Naive Vorstellungen	KV-NA-UK: Unkonkret/Keine Vorstellung	11	3	11	3	N=0
		KV-NA-FV: Fehlvorstellungen	1	0	1	0	
		Summe NA			**12**	**3**	
		Einzelzählung NA	**12**	**3**			
	KV-TG: Teilbereich der Grunddefinition	KV-TG-DF: Denk-und Wahrnehmungsfehler	3	6	3	6	N=1
		KV-TG-IF: Einfluss von internen Faktoren	1	7	1	8	
		KV-TG-EF: Einfluss von externen Faktoren	0	2	0	2	
	Summe TG				**4**	**16**	
	Einzelzählung TG		**4**	**15**			

Im Gegensatz dazu zeigte sich ein deutlich verbessertes Verständnis kognitiver Verzerrungen im Post-Test, in dem nur noch drei Personen Schwierigkeiten mit der Definition aufwiesen (KV-NA-UK). 81,3 % (13 von 16) erreichten die höchste Niveaustufe 1, wobei sechs Proband*innen kognitive Verzerrungen als *Wahrnehmungsfehler* (KV-TG-DF) interpretierten und acht Teilnehmende darunter *interne Einflüsse* (KV-TG-IF) auf Denk- und Bewertungsprozesse fassten. *Externe Faktoren* (KV-TG-EF) wurden indes von zwei Befragten als Bestandteil kognitiver Verzerrungen erkannt.

Die Antworten der offenen Fragen VF3 und EF1 wurden bezüglich der angeführten *Bewertungsfaktoren* (BF) ausgewertet und entsprechend gruppiert (Tabelle 84). Auf Ebene der *selbstreflexiven Aspekte* (BF-SR) zeigte sich im Pre-Test, dass vor allem das *Fach- oder Vorwissen* (BF-SR-FW) für den größten Anteil der Befragten (68,6 %, 11 von 16) als relevanter Faktor angesehen wurde. Weitere *interne Faktoren* (BF-SR-IF), wie *Emotionen* (BF-SR-IF-EM), *persönliches Interesse* (BF-SR-IF-IN) oder die eigene *Befangenheit* (BF-SR-IF-BV), fanden jeweils bei zwei Teilnehmenden Erwähnung. Unter den *externen Faktoren* (BF-SR-EX) wurde vor allem die *soziale Norm* (BF-SR-SN) als größter Einflussfaktor angeführt (37,5 %, 6 von 16), aber auch die *Bildung* (BF-SR-

BI) von fünf Personen oder die *Medien* (BF-SR-ME) von vier Personen als beeinflussend wahrgenommen. Insgesamt wurden im Schnitt pro Proband*in 2,25 verschiedene selbstreflexive Faktoren im Pre-Test benannt.

Tabelle 84: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *selbstreflexive Aspekte* (BF-RA), *N* = 16

Kat.	1.UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
BF: Bewertungsfaktoren	SR: Selbstreflexive Faktoren	SR-IF: Interne Faktoren	BF-SR-IF-FW: Fachwissen	11	10	16	14
			BF-SR-IF-PR: Persönliche Relevanz	1	2	1	3
			BF-SR-IF-MW: Moral und Werte	1	3	1	3
			BF-SR-IF-MW: Handlungen	0	0	0	0
			BF-SR-IF-EM: Emotionen	2	2	2	3
			BF-SR-IF-EF: Erfahrungen	1	6	1	8
			BF-SR-IF-IN: Interesse	2	3	2	3
			BF-SR-IF-BV: Befangenheit, Vorurteile	2	5	3	7
			BF-SR-IF-PK: Persönlichkeit	0	2	0	3
			BF-SR-IF-KV: Kognitive Verzerrungen	0	1	0	1
			Summe SR-IF			**26**	**45**
			Einzelzählung SR-IF	**20**	**34**		
		SR-EF: Externe Faktoren	BF-SR-EF-SN: Soziale Norm	6	9	7	12
			BF-SR-EF-ME: Medien	4	7	4	9
			BF-SR-EF-BI: Studium/Bildung	5	4	5	5
			BF-SR-EF-PO: Politische Position	0	1	0	1
			BF-SR-EF-RL: Religion	1	0	1	0
			Summe SR-EF			**17**	**27**
			Einzelzählung SR-EF	**16**	**21**		
			Summe SR Gesamt			**43**	**72**
			Einzelzählung SR Gesamt	**36**	**55**		

Im Post-Fragebogen stieg die durchschnittliche Anzahl differenzierter selbstreflexiver Aspekte pro Proband*in auf 3,4 und auch der wahrgenommene Einfluss interner und externer Faktoren nahm zu. So wurde unter anderem der zuvor nur von einer Person berücksichtigte Aspekt der gemachten *Erfahrungen* (BF-SR-IF-EF) im Post-

Test von sechs Teilnehmenden (37,5 %) aufgegriffen. Ebenso erhöhte sich die Aufmerksamkeit für den Einfluss der eigenen *Befangenheit* (BF-SR-IG-BV) von 12,5 % (2 von 16) auf 31,3 % (5 von 16). Die Sensibilität für die *soziale Norm* (BF-SR-EF-SN) stieg auf 56,3 % (9 von 16) und das Bewusstsein für den Einfluss von *Medien* (BF-SR-EF-ME) wurde im Post-Test von 43,8 % der Teilnehmenden als relevanter Einflussfaktor reflektiert.

Im Pre-Test wurden lediglich von sieben der 16 Befragten (43,8 %) Bewertungsfaktoren mit *epistemologischem Bezug* (BF-EP) angeführt, wodurch der Mittelwert differenter Aspekte pro Person nur bei 0,4 lag. Lediglich die *Autorität von Expert*innen* (BF-EP-EX) (3 von 16, 18,8 %) und die *Herkunft einer Information* (BF-EP-HK) (2 von 16, 12,5 %) wurden von mehr als einer Person als relevante Einflussfaktoren bei der Bewertung benannt (Tabelle 85).

Tabelle 85: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *epistemologische Aspekte* (BF-EP), *N* = 16

Kat.	1.UK 2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
BF: Bewertungsfaktoren	EP: Epistemologische Faktoren	BF-EP-EX: Autoritäten/Expertise	3	2	3	3
		BF-EP-AK: Aktualität	0	2	0	2
		BF-EP-IK: Interessenskonflikte	0	0	0	0
		BF-EP-PR: Präsentation	1	2	3	4
		BF-EP-HK: Herkunft der Information	2	7	4	8
		BF-EP-ME: Methodik	1	0	1	0
		BF-EP-KO: Konsistenzprüfung	0	0	0	0
		BF-EP-TR: Transparenz	0	0	0	0
		BF-EF-SE: Soziokulturelle Eingebundenheit	0	3	0	3
		Summe EP			**11**	**20**
		Einzelzählung EP	**7**	**16**		

Im Post-Test nahm die Aufmerksamkeit für Einflussfaktoren mit epistemologischem Bezug leicht zu, wobei 12 der 16 Teilnehmenden (75 %) entsprechende Aspekte anführten. Sieben Personen bezeichneten die *Herkunft einer Information* (BF-EP-HK) als relevanten Einflussfaktor und die *soziokulturelle Eingebundenheit* (BF-EP-SE)

wurde von drei Teilnehmenden berücksichtigt. Die durchschnittliche Anzahl differenzierter Aspekte pro Person stieg geringfügig auf 1,0.

Wissensbezogene Aspekte (BF-WI) wurden von den Studierenden trotz der gezielten Fragestellung nach Einflussfaktoren bei der Bewertung von Themen mit naturwissenschaftlichem Bezug nur selten genannt (Tabelle 86). Die Ergebnisse von Pre- und Post-Test blieben auf vergleichbar niedrigem Niveau und zeigten keine signifikanten Unterschiede. Der am häufigsten genannte Aspekt war die Einhaltung *wissenschaftlicher Standards* (BF-WI-WS), die jedoch meist wenig konkret beschrieben wurden.

Tabelle 86: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *naturwissenschaftliche Aspekte* (BF-NW), $N = 16$

Kat.	1.UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
BF: Bewertungsfaktoren	NW: Naturwissenschaftliche Faktoren		BF-NW-WS: Wissenschaftliche Standards	3	4	3	5
			BF-NW-EF: Effizenz	0	0	0	0
			BF-NW-KA: Kostenabwägung	0	0	0	0
			BF-NW-NA: Auswirkungen & Nachhaltigkeit	1	1	1	1
			BF-NW-FT: Funktionsweise	0	0	0	0
			BF-NW-UM: Umsetzbarkeit	0	0	0	0
			BF-NW-RH: Rohstoffherkunft	0	0	0	0
			BF-NW-NO: Notwendigkeit	0	0	0	0
		Summe NW				**4**	**6**
		Einzelzählung NW		**4**	**5**		

Der geringe Bezug auf naturwissenschaftliche Aspekte könnte auf die geringere Kontextualisierung des Fragebogenformats im Vergleich zur Interviewerhebung zurückzuführen sein, was zu weniger spezifischen und weniger reflektierten Antworten führte.

Die Überprüfung aller Ergebnisse der offenen Fragestellungen im Vergleich zwischen Pre- und Post-Test zeigte signifikante Verbesserungen im *Verständnis der Selbstreflexion* (VS) mit $z = -2{,}511$, $p = 0{,}014$ und einer mittleren Effektstärke von $r = 0{,}444$ (Tabelle 87). Das *Verständnis Kritischen Denkens* (VK) verbesserte sich eben-

falls signifikant und erreichte mit z = -2,873 und p = 0,002 sogar eine starke Effektstärke von r = 0,508. Unter den Bewertungsfaktoren zeigte sich der Unterschied im Bereich *wissensbezogener Aspekte* (BF-WI) als nicht signifikant (z = -0,577, p = 1,000, r = 0,102), wohingegen die Werte der *epistemologischen Aspekte* (BF-EP) die Signifikanzgrenze knapp überschritten (z = -2,165, p = 0,047, r = 0,383).

Tabelle 87: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der offenen Fragestellungen, N = 16

S3.F1&2 - Wilcoxon-Test

Teststatistiken[a]	VS_Post - VS_Pre	VK_Post - VK_Pre	BFWI_Post - BFWi_Pre	BFEP_Post - BFEP_Pre	BFSR_Post - BFSR_Pre	BFSRIF_Post - BFSRIF_Pre	BFSREF_Post - BFSREF_Pre	KV_Post - KV_Pre
Z	-2,511b	-2,873b	-,577b	-2,165b	-2,428b	-2,435b	-1,184b	-2,653b
Asymp. Sig. (2-seitig)	0,012	0,004	0,564	0,030	0,015	0,015	0,236	0,008
Exakte Sig. (2-seitig)	0,014	0,002	1,000	0,047	0,016	0,018	0,340	0,010
Exakte Sig. (1-seitig)	0,007	0,001	0,500	0,023	0,008	0,009	0,170	0,005
Punkt-Wahrscheinlichkeit	0,006	0,001	0,375	0,020	0,005	0,007	0,035	0,004
a Wilcoxon-Test								
b Die Summe der negativen Ränge ist gleich der Summe der positiven Ränge.								
c Basiert auf negativen Rängen.								
Effektstärke	0,444	0,508	0,102	0,383	0,429	0,430	0,209	0,469

Eine deutlich signifikantere Perspektiverweiterung war bei den *selbstreflexiven Aspekten* (BF-SR) zu verzeichnen (z = -2,428, p = 0,016, r = 0,429), welche sich schwerpunktmäßig auf die Reflexion *interner Faktoren* (BF-SR-IF) zurückführen ließ (z = -2,435, p = 0,018, r = 0,430). Die Unterschiede bei der Reflexion *externer Faktoren* (BF-SR-EF) verblieben nicht signifikant (z = -1,184, p = 0,340, r = 0,209). Eine größere Veränderung war hingegen bezüglich des *Verständnisses kognitiver Verzerrungen* (KV) erkenntlich, welches sich mit z = -2, 653 und p = 0,010 bei einer mittleren Effektstärke von r = 0,469 signifikant verbesserte.

Die Einschätzungen der Teilnehmenden zu den geschlossenen Fragen zur *Introspektion* (IS) zeigten sich im Pre-Test weitgehend gleichbleibend auf einem ähnlichen Niveau und lagen durchschnittlich bei 3,363 auf der fünfstufigen Likert-Skala (Tabelle 88). Im Post-Test gaben die Teilnehmenden in allen Fällen höhere Zustimmungswerte an, insbesondere bei der *Reflexion persönlicher Erfahrungen und Hintergründe* (IS3) und der *Reflexion von Vorurteilen und -annahmen* (IS5). Dies führte zu einer Erhöhung des Durchschnittswertes auf 3,950. Das Minimum der Zustimmung über alle fünf Items und Teilnehmenden stieg von 1 auf 2, während der Median bei allen Teilnehmenden auf 4 anstieg. Die Varianzen der Pre- als auch der Post-Tests der Items IS1

bis IS5 zeigten für eine fünfstufige Likert-Skala geringe Abweichungen vom Mittelwert mit Werten unterhalb von 1,0 (vgl. Kap. 4.2.3), was auf eine gewisse Homogenität der Antworten schließen lässt. Diese Ergebnisse und beobachteten Veränderungen deuten damit auf eine subjektiv empfundene erhöhte Sensibilität für Einflussfaktoren und damit verbundene reflexive Prozesse im Post-Test hin.

Tabelle 88: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Introspektion* (IS), *N* = 16

Pre							
Code	**IS1**	**IS2**	**IS3**	**IS4**	**IS5**	**Summe IS**	**Summe rel. %**
Summe	53	53	54	55	54	269	1076,0
Mittelwert	**3,313**	**3,313**	**3,375**	**3,438**	**3,375**	16,813	67,3
Stabw.	0,793	0,602	0,957	0,892	0,619	2,786	11,1
Varianz	0,629	0,363	0,917	0,796	0,383	7,763	124,2
Modus	4	3	4	4	3	20,0	80,0
Median	3,5	3	3,5	4	3	17,0	68,0
Min	2	2	1	2	2	12,0	48,0
Max	4	4	5	5	4	20,0	80,0

Post							
Code	**IS1**	**IS2**	**IS3**	**IS4**	**IS5**	**Summe IS**	**Summe rel. %**
Summe	62	63	66	61	64	316	1264,0
Mittelwert	**3,875**	**3,938**	**4,125**	**3,813**	**4,000**	19,750	79,0
Stabw.	0,619	0,772	0,500	0,544	0,730	2,082	8,3
Varianz	0,383	0,596	0,250	0,296	0,533	4,333	69,3
Modus	4	4	4	4	4	20,0	80,0
Median	4	4	4	4	4	20,0	80,0
Min	3	2	3	3	3	17,0	68,0
Max	5	5	5	5	5	25,0	100,0

IS1:	Wie oft machen Sie sich Ihre eigenen Gedanken und Gefühle bewusst, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?
IS2:	Inwieweit setzen Sie sich mit deinen eigenen Moralvorstellungen und Werte auseinander, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?
IS3:	Wie häufig denken Sie über Ihre persönlichen Erfahrungen und Hintergründe nach, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?
IS4:	Inwiefern reflektieren Sie Ihre Interessen und Motivationen, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?
IS5:	Wie häufig hinterfragen Sie Ihre eigenen Vorurteile und Annahmen, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?

Die Veränderungen der Items des Blocks *Introspektion* (IS) zwischen Pre- und Post-Test fielen bei allen Items bis auf IS4 (*Interessen und Motivationen*) signifikant bei mittlerer Effektstärke aus (Tabelle 89).

Tabelle 89: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Blocks *Introspektion* (IS), *N* = 16

Frage	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
IS1	0,035	-2,309b	0,408	Signifikant
IS2	0,018	-2,486b	0,439	Signifikant
IS3	0,028	-2,292b	0,405	Signifikant
IS4	0,186	-1,513b	0,267	Nicht signifikant
IS5	0,036	-2,233b	0,395	Signifikant

Die Ergebnisse des Items EF2, das nach der Fähigkeit zur rein fachlichen Bewertung in Bewertungsprozessen fragte, ergaben im Pre-Test einen Mittelwert von 3,250 auf der Likert-Skala, was auf eine mäßige Einschätzung der eigenen Fähigkeit zur rein fachlichen Bewertung hinweist (Tabelle 90).

Tabelle 90: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Einflussfaktoren 2 (EF2), *N* = 16

Code	**EF2 Pre**	**EF2 Post**	**EF2 rel.Pre**	**EF2 rel.Post**
Summe	52	60	1040,0	1200,0
Mittelwert	**3,250**	**3,750**	65,0	75,0
Stabw.	0,775	0,447	15,5	8,9
Varianz	0,600	0,200	240,0	80,0
Modus	4	4	80	80
Median	3	4	60	80
Min	2	3	40	60
Max	4	4	80	80

EF2:	Inwieweit glauben Sie, dass Sie in der Lage sind, vornehmlich auf fachlicher Ebene zu urteilen?

Dieser Wert stieg im Post-Test signifikant (z = -2,309, p = 0,035, r = 0,408) auf 3,750, was eine Verbesserung des persönlichen Empfindens, auf rein fachlicher Ebene urteilen zu können, ausdrückt (Tabelle 91). Hierbei sei jedoch noch einmal betont, dass dieses Item die subjektive und nicht-kontextualisierte Selbsteinschätzung der Fähigkeit wiedergibt, rein auf fachlicher Ebene bewerten zu können. Dieses Empfinden kann von der tatsächlichen Urteilsfähigkeit abweichen und spiegelt nur die persönliche Überzeugung wider. Dennoch lässt sich ein gesteigertes Selbstbewusstsein hinsichtlich des eigenen Bewertungsvermögens erkennen, was möglicherweise auf die explizite Thematisierung von Strategien zur Bewertung und Positionierung im Rahmen der Vermittlungseinheit zurückzuführen ist.

Tabelle 91: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Items EF2, *N* = 16

Frage	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
EF2	0,035	-2,309b	0,408	Signifikant

Die Ergebnisse der personenbezogenen Items der Fragestellung EF4 (Tabelle 92) zeigten im Pre-Test eine subjektiv wahrgenommene erhöhte Reflexionsbereitschaft für den Einfluss der *eigenen Grundeinstellung* (EF4_4), der eigenen *Wert- und Moralvorstellungen* (EF4_3) sowie für das *eigene Wissen*, was sich in Mittelwerten über 3,500 ausdrückte (EF4_12). Deutlich weniger häufig wurde im Alltag die *eigene Religion* (EF4_9) oder das *eigene Geschlecht* (EF4_7) reflektiert.

Tabelle 92: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Einflussfaktoren 4* (EF4), personenbezogene Einflüsse, *N* = 16

Pre																
Code	**EF4 1**	**EF4 2**	**EF4 3**	**EF4 4**	**EF4 5**	**EF4 6**	**EF4 7**	**EF4 8**	**EF4 9**	**EF4 10**	**EF4 11**	**EF4 12**	**EF4 13**	**EF4 14**	**EF4 15**	**EF4 16**
Summe	53	47	59	65	56	39	34	43	27	36	44	59	53	49	56	44
MW	**3,313**	**2,938**	**3,688**	**4,063**	**3,500**	**2,438**	**2,125**	**2,688**	**1,688**	**2,250**	**2,750**	**3,688**	**3,313**	**3,063**	**3,500**	**2,750**
Stabw.	1,078	0,929	0,704	0,443	0,816	0,892	1,258	1,014	1,014	1,000	0,931	0,704	0,793	1,237	0,966	1,238
Varianz	1,163	0,863	0,496	0,196	0,667	0,796	1,583	1,029	1,029	1,000	0,867	0,496	0,629	1,529	0,933	1,533
Modus	4	3	4	4	4	2	2	3	1	2	3	4	4	4	4	3
Median	4	3	4	4	4	2	2	3	1	2	3	4	3,5	3	4	3
Min	1	2	2	3	2	1	1	1	1	1	1	2	2	1	2	1
Max	5	5	5	5	5	4	5	4	4	4	4	5	4	5	5	5

Post																
Code	**EF4 1**	**EF4 2**	**EF4 3**	**EF4 4**	**EF4 5**	**EF4 6**	**EF4 7**	**EF4 8**	**EF4 9**	**EF4 10**	**EF4 11**	**EF4 12**	**EF4 13**	**EF4 14**	**EF4 15**	**EF4 16**
Summe	57	59	68	65	60	45	43	55	34	43	55	70	65	62	67	61
MW	**3,563**	**3,688**	**4,250**	**4,063**	**3,750**	**2,813**	**2,688**	**3,438**	**2,125**	**2,688**	**3,438**	**4,375**	**4,063**	**3,875**	**4,188**	**3,813**
Stabw.	0,892	0,873	0,577	0,574	0,577	1,167	1,302	0,964	1,408	0,946	0,727	0,500	0,680	0,806	0,544	1,167
Varianz	0,796	0,763	0,333	0,329	0,333	1,363	1,696	0,929	1,983	0,896	0,529	0,250	0,463	0,650	0,296	1,363
Modus	4	4	4	4	4	2	2	3	1	3	3	4	4	4	4	4
Median	4	4	4	4	4	2,5	2,5	3	1,5	3	3	4	4	4	4	4
Min	1	2	3	3	3	1	1	1	1	1	2	4	3	2	3	1
Max	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5

EF4:	Wie sehr hinterfragen Sie folgende Faktoren, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	
Personenbezogene Einflüsse	EF4_1	Eigene Persönlichkeit
	EF4_2	Stimmung und Emotion
	EF4_3	Werte und Moralvorstellungen
	EF4_4	Eigene Grundeinstellung
	EF4_5	Eigene Bildung
	EF4_6	Eigenes Alter
	EF4_7	Eigenes Geschlecht
	EF4_8	Eigene soziale Schicht
	EF4_9	Eigene Religion
	EF4_10	Eigene Ethnie
	EF4_11	Eigene Kultur
	EF4_12	Wissen
	EF4_13	Medialer Kontext
	EF4_14	Politische Einstellung
	EF4_15	Erfahrungen und Erinnerungen
	EF4_16	Selbstkonzept und Selbstwertgefühl

Alle personenbezogenen Items zeigten im Post-Test erhöhte Werte, wobei insbesondere das *Selbstkonzept und Selbstwertgefühl* (EF4_16), die *politische Einstellung* (EF4_14) sowie die *Stimmung und Emotion* (EF4_2) deutlich höhere Werte aufwiesen und im Mittel um mindestens 0,750 stiegen. Aber auch die *eigene soziale Schicht* (EF4_8) und der *mediale Kontext* (EF4_13) wurden deutlich häufiger reflektiert.

Unter den Einflüssen mit Bezug zu kognitiven Verzerrungen fanden im Pre-Test die *eigene Motivation* (EF4_18) und die *Komplexität von Informationen* (EF4_26), mit Mittelwerten knapp über 3,200, vermehrt Beachtung. Am wenigsten wurden *Gruppendynamiken* (EF4_21) und *Übergeneralisierungen* (EF4_25) bei alltäglichen Bewertungen berücksichtigt (Tabelle 93).

Tabelle 93: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Einflussfaktoren 4* (EF4), kognitive Verzerrungen, $N = 16$

Pre												
Code	**EF4 17**	**EF4 18**	**EF4 19**	**EF4 20**	**EF4 21**	**EF4 22**	**EF4 23**	**EF4 24**	**EF4 25**	**EF4 26**	**Summe EF4**	**Summe EF4 rel.**
Summe	51	54	51	51	43	49	51	47	43	52	1256	966,2
MW	**3,188**	**3,375**	**3,188**	**3,188**	**2,688**	**3,063**	**3,188**	**2,938**	**2,688**	**3,250**	78,500	60,4
Stabw.	1,047	1,025	0,911	0,981	1,014	1,063	0,834	0,929	0,873	0,775	12,160	9,4
Varianz	1,096	1,050	0,829	0,963	1,029	1,129	0,696	0,863	0,763	0,600	147,867	87,5
Modus	4	4	3	3	2	3	3	3	2	4	80,0	61,5
Median	3	3,5	3	3	3	3	3	3	3	3	79,5	61,2
Min	1	2	2	2	1	1	2	1	1	2	60,0	46,2
Max	5	5	5	5	4	5	5	4	4	4	102,0	78,5

Post												
Code	**EF4 17**	**EF4 18**	**EF4 19**	**EF4 20**	**EF4 21**	**EF4 22**	**EF4 23**	**EF4 24**	**EF4 25**	**EF4 26**	**Summe EF4**	**Summe EF4 rel.**
Summe	64	63	58	57	55	59	59	59	55	59	1497	1151,5
MW	**4,000**	**3,938**	**3,625**	**3,563**	**3,438**	**3,688**	**3,688**	**3,688**	**3,438**	**3,688**	93,563	72,0
Stabw.	0,632	0,680	0,957	1,031	1,031	0,602	0,873	0,793	0,964	0,793	9,633	7,4
Varianz	0,400	0,463	0,917	1,063	1,063	0,363	0,763	0,629	0,929	0,629	92,796	54,9
Modus	4	4	4	3	3	4	3	4	3	4	92,0	70,8
Median	4	4	4	3	3	4	4	4	3	4	92,0	70,8
Min	3	3	2	2	2	3	2	2	2	2	75,0	57,7
Max	5	5	5	5	5	5	5	5	5	5	109,0	83,8

EF4:		Wie sehr hinterfragen Sie folgende Faktoren, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?
Kognitive Verzerrungen	EF4_17	Soziale Normen und Erwartungen
	EF4_18	Motivation und Bedürfnisse
	EF4_19	Verfügbarkeit von Alternativen
	EF4_20	Intuition und Bauchgefühl
	EF4_21	Gruppendynamik und Konformitätsdruck
	EF4_22	Erwartungen und Vorurteile
	EF4_23	Framing und Präsentation von Information
	EF4_24	Selektive Wahrnehmung und Bestätigungsfehler
	EF4_25	Übergeneralisierung und Kategorisierung
	EF4_26	Komplexität der Information

Im Post-Test waren ebenfalls sämtliche Faktoren mit Bezug zu kognitiven Verzerrungen klar erhöht. Insbesondere die Sensibilität für den Einfluss *sozialer Normen* (EF4_17), von *Gruppendynamiken* (EF4_21), *selektiver Wahrnehmung* (EF4_24) sowie *Übergeneralisierung* (EF4_25) stieg im Mittel um mindestens 0,750 an. Die Varianzen des Großteils der Items von EF4 ließen nur niedrige bis moderate Abweichungen

von den errechneten Mittelwerten erkennen. Die größte Heterogenität der Antworten wurde bei EF4_9 (*Religion*) und EF4_7 (*Geschlecht*) mit Werten von über 1,600 deutlich. Dies deutet auf unterschiedlich starke Sensibilitäten in diesen Bereichen hin, die möglicherweise durch persönliche Erfahrungen oder Überzeugungen beeinflusst wurden.

Tabelle 94: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der Items zu EF4, *N* = 16

Frage	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
EF4_1	0,344	-1,100b	0,194	Nicht signifikant
EF4_2	0,028	-2,292b	0,405	Signifikant
EF4_3	0,008	-2,714b	0,480	Signifikant
EF4_4	1,000	,000c	0,000	Nicht signifikant
EF4_5	0,359	-1,265b	0,224	Nicht signifikant
EF4_6	0,109	-1,897b	0,335	Nicht signifikant
EF4_7	0,047	-2,165b	0,383	Signifikant
EF4_8	0,028	-2,292b	0,405	Signifikant
EF4_9	0,219	-1,382b	0,244	Nicht signifikant
EF4_10	0,305	-1,125b	0,199	Nicht signifikant
EF4_11	0,057	-2,066b	0,365	Nicht signifikant
EF4_12	0,004	-2,810b	0,497	Signifikant
EF4_13	0,022	-2,392b	0,423	Signifikant
EF4_14	0,009	-2,648b	0,468	Signifikant
EF4_15	0,027	-2,326b	0,411	Signifikant
EF4_16	0,044	-2,034b	0,360	Signifikant
EF4_17	0,016	-2,410b	0,426	Signifikant
EF4_18	0,057	-2,070b	0,366	Nicht signifikant
EF4_19	0,254	-1,308b	0,231	Nicht signifikant
EF4_20	0,183	-1,604b	0,284	Nicht signifikant
EF4_21	0,009	-2,652b	0,469	Signifikant
EF4_22	0,027	-2,352b	0,416	Signifikant
EF4_23	0,035	-2,309b	0,408	Signifikant
EF4_24	0,009	-2,652b	0,469	Signifikant
EF4_25	0,018	-2,443b	0,432	Signifikant
EF4_26	0,092	-1,941b	0,343	Nicht signifikant

Die Ergebnisse des Wilcoxon-Tests für die Items der *Einflussfaktoren* (EF) (Tabelle 94) zeigten signifikante Unterschiede mit mittlerer Effektstärke innerhalb der personenbezogenen Einflüsse in den Bereichen *Stimmung und Emotion* (EF4_2), *Werte- und Moralvorstellungen* (EF4_3), *Geschlecht* (EF4_7), *soziale Schicht* (EF4_8), *Wissen* (EF4_12), *medialer Kontext* (EF4_13), *politische Einstellung* (EF4_14), *Erfahrungen* (EF4_15) und *Selbstkonzept* (EF4_16). Unter den Einflüssen mit Bezug zu kognitiven Verzerrungen wiesen die Faktoren *soziale Norm* (EF4_17), *Gruppendynamik* (EF4_21),

Erwartungen und Vorurteilen (EF4_22), *Präsentation von Informationen* (EF4_ 23), *selektive Wahrnehmung* (EF4_24) und *Übergeneralisierung* (EF4_25) ebenfalls signifikante Ergebnisse bei mittlerer Effektstärke auf.

Im Bereich *Bewertung und Gewichtung* (BG) (Tabelle 95) tendierten die Befragten im Pre-Test zu einer eher schwächeren *Einschätzung* (BG1) und *Berücksichtigung* (BG2) von Einflussfaktoren, was die Mittelwerte von 2,875 nahelegen.

Tabelle 95: S3.F1&2 - Ergebnisse der Fragebogenerhebung, *Bewertung und Gewichtung* (BG), *N* = 16

Pre				
Code	**BG1**	**BG2**	**Summe BG**	**Summe BG rel. %**
Summe	46	46	92	920,0
Mittelwert	**2,875**	**2,875**	5,750	57,5
Stabw.	0,957	1,025	1,483	14,8
Varianz	0,917	1,050	2,200	220,0
Modus	4	3	7,0	70,0
Median	3	3	6,0	60,0
Min	1	1	2,0	20,0
Max	4	5	8,0	80,0

Post				
Code	**BG1**	**BG2**	**Summe BG**	**Summe BG rel. %**
Summe	58	57	115	1150,0
Mittelwert	**3,625**	**3,5625**	7,188	71,9
Stabw.	0,806	0,814	1,223	12,2
Varianz	0,650	0,663	1,496	149,6
Modus	4	3	7,0	70,0
Median	4	3,5	7,0	70,0
Min	2	2	5,0	50,0
Max	5	5	9,0	90,0

BG1:	Wie stark schätzen sie den Einfluss der vorhin genannten Faktoren (Bauchgefühl, Stimmung usw.) im Vergleich zum Fachwissen, wenn Sie ein Thema mit naturwissenschaftlichem Bezug bewerten sollen?
BG2:	Inwieweit wägen Sie diese Faktoren (Bauchgefühl. Stimmung usw.) bewusst gegen Ihr Fachwissen ab, wenn sie ein Thema mit naturwissenschaftlichem Bezug bewerten sollen?

Diese Mittelwerte stiegen im Post-Test deutlich auf Werte über 3,500 an, bei gleichzeitig geringen Abweichungen vom Mittelwert, was auf eine subjektiv empfundene höhere Aufmerksamkeit für Einflüsse in Bewertungssituationen hinweist.

Beide Items ließen im Vergleich zwischen Pre- und Post-Test signifikante Unterschiede bei mittlerer Effektstärke erkennen (Tabelle 96).

Tabelle 96: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Blocks *Bewertung und Gewichtung (BG)*, *N* = 16

Frage	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
BG1	0,004	-2,739b	0,484	Signifikant
BG2	0,043	-2,153b	0,381	Signifikant

Der SRIS zeigte im Pre-Test auf der Subskala der *Selbstreflexion* (SR) einen Mittelwert von 5,094 bei geringer Streuung über alle Items und Proband*innen hinweg (Tabelle 97). Im Gegensatz dazu lag der Mittelwert auf der Subskala der *Einsicht* (IN) mit 4,646 darunter und wies eine deutlich größere Streuung auf, was auf eine Heterogenität in der Fähigkeit hindeutet, aus selbstreflexivem Verhalten Einsichten gewinnen zu können.

Tabelle 97: S3.F1&2 - Ergebnisse der Fragebogenerhebung, SRIS nach Silvia (2021), *N* = 16

Pre								
Code	**SR3**	**SR5**	**SR6**	**SR8**	**SR9**	**SR10**	**Summe SR**	**Summe SR rel.**
Summe	85	82	79	74	85	84	489	1164,3
Mittelwert	**5,313**	**5,125**	**4,938**	**4,625**	**5,313**	**5,250**	30,563	72,8
Stn. Abw.	1,195	1,088	1,289	1,088	1,078	1,000	5,020	12,0
Varianz	1,429	1,183	1,663	1,183	1,163	1,000	25,196	142,8
Modus	6	6	5	5	5	5	33,0	78,6
Median	6	5	5	5	5	5	31,5	75,0
Min	3	3	2	2	3	3	20,0	47,6
Max	7	7	7	6	7	7	37,0	88,1

Post								
Code	**SR3**	**SR5**	**SR6**	**SR8**	**SR9**	**SR10**	**Summe SR**	**Summe SR rel.**
Summe	93	95	94	92	104	93	571	1359,5
Mittelwert	**5,813**	**5,938**	**5,875**	**5,750**	**6,500**	**5,813**	35,688	85,0
Stn. Abw.	1,109	0,680	1,147	0,931	0,632	1,223	4,453	10,6
Varianz	1,229	0,463	1,317	0,867	0,400	1,496	19,829	112,4
Modus	7	6	7	6	7	7	38,0	90,5
Median	6	6	6	6	7	6	36,5	86,9
Min	4	5	4	4	5	3	26,0	61,9
Max	7	7	7	7	7	7	42,0	100,0

Pre										
Code	**IN2**	**IN4**	**IN5**	**IN6**	**IN7**	**IN8**	**Summe IN**	**Summe IN rel.**	**Summe SR+IN Gesamt**	**Summe SR+IN Ges. rel.**
Summe	71	74	83	64	75	79	446	1061,9	935	1113,1
Mittelwert	**4,438**	**4,625**	**5,188**	**4,000**	**4,688**	**4,938**	27,875	66,4	58,438	69,6
Stn. Abw.	1,750	1,544	1,424	1,461	1,740	1,436	6,365	15,2	7,882	9,4
Varianz	3,063	2,383	2,029	2,133	3,029	2,063	40,517	229,7	62,129	88,1
Modus	5	6	6	5	5	6	32,0	76,2	62,0	73,8
Median	5	5	6	4	5	5,5	29,5	70,2	59,0	70,2
Min	1	2	1	2	1	2	15,0	35,7	45,0	53,6
Max	7	7	7	6	7	7	37,0	88,1	71,0	84,5

Post										
Code	**IN2**	**IN4**	**IN5**	**IN6**	**IN7**	**IN8**	**Summe IN**	**Summe IN rel.**	**Summe Gesamt**	**Summe Ges. rel.**
Summe	80	84	91	75	83	89	502	1195,2	1073	1277,4
Mittelwert	**5,000**	**5,250**	**5,688**	**4,688**	**5,188**	**5,563**	31,375	74,7	67,063	79,8
Stn. Abw.	1,414	1,183	1,195	1,302	1,471	0,629	3,981	9,5	6,148	7,3
Varianz	2,000	1,400	1,429	1,696	2,163	0,396	15,850	89,9	37,796	53,6
Modus	5	6	6	5	6	6	33,0	78,6	73,0	86,9
Median	5	6	6	5	6	6	33,0	78,6	68,0	81,0
Min	3	3	2	3	2	4	22,0	52,4	56,0	66,7
Max	7	7	7	7	7	6	38,0	90,5	78,0	92,9

Pre	**Mittelwert SR Pre**	**Mittelwert IN Pre**	**Mittelwert Gesamt Pre**
	5,094	**4,646**	**4,870**
Post	**Mittelwert SR Post**	**Mittelwert IN Post**	**Mittelwert Gesamt Post**
	5,948	**5,229**	**5,589**

Im Post-Test wiesen die Werte der Items im Mittel eine deutliche Erhöhung auf und auch die Streuung um die Mittelwerte nahm ab. Der Durchschnittswert der Subskala der *Selbstreflexion* (SR) stieg auf 5,948, während sich der Wert der Subskala der *Einsicht* (IN) auf 5,229 erhöhte. Daraus lässt sich ableiten, dass die Studierenden im Post-Test eine subjektiv gesteigerte Tendenz zu selbstreflexivem Verhalten empfanden und auch die damit verbundenen Einsichten besser einordnen konnten.

Alle Items der Subskala *Selbstreflexion* (SR) waren mit mittlerer bis hoher Effektstärke signifikant erhöht (Tabelle 98). Im Gegensatz dazu zeigten die Items der Subskala *Einsicht* (IN) zwar Erhöhungen der Mittelwerte, diese waren aber bis auf Item IN6 nicht signifikant.

Tabelle 98: S3.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der SRIS, *N* = 16

Frage	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
SR3	0,021	-2,530b	0,447	Signifikant
SR5	0,002	-2,919b	0,516	Signifikant
SR6	0,004	-2,830b	0,500	Signifikant
SR8	<,001	-3,140b	0,555	Signifikant
SR9	<,001	-3,134b	0,554	Signifikant
SR10	0,045	-2,179b	0,385	Signifikant
IN2	0,193	-1,410b	0,249	Nicht signifikant
IN4	0,156	-1,569b	0,277	Nicht signifikant
IN5	0,055	-2,126b	0,376	Nicht signifikant
IN6	0,039	-2,209b	0,390	Signifikant
IN7	0,201	-1,417b	0,250	Nicht signifikant
IN8	0,097	-1,705b	0,301	Nicht signifikant

Als letzter hier dargestellter Block wurde der CRT ausschließlich im Pre-Test durchgeführt, da die Kenntnis der Fragen eine valide Durchführung im Post-Test unmöglich gemacht hätte. Zudem wurden die Teilnehmenden, die bereits an der Fragebogenstudie S2.F1 mitgewirkt hatten, von der Auswertung ausgeschlossen, sodass sich die Zahl der vollständigen Datensätze auf *N* = 14 belief.

Tabelle 99: S3.F1&2 - Ergebnisse der Fragebogenerhebung, CRT nach Frederick (2005), *N* = 14

Code	**CRT1**	**CRT2**	**CRT3**	**Korrekt**	**%**
Korrekt	3	7	9	19	**45,2**

Verteilung der Ergebisse in %			
28,6	21,4	35,7	14,3
0	1	2	3
Low			High

Die Ergebnisse machen eine ähnliche Verteilung wie in den Studien S2.F1 und S1.F3 mit einem Mittelwert von 1,357 korrekten Antworten pro Person sichtbar. Nur 14,3 % (2 von 14) der Befragten beantworteten alle Fragen korrekt, während 28,6 % (4 von 14) keine der Fragen lösten (Tabelle 99). Die Verteilung lag im erwartbaren Rahmen und lässt auf eine verhältnismäßig starke Einwirkung des intuitiven Systems 1 auf die Beantwortung der Fragen schließen.

Zusammenfassend ließen sich in der Fragebogenerhebung S3.F1&2 signifikante Veränderungen in den Antworttendenzen der einzelnen Blöcke erkennen. Sowohl die Tendenzen, selbstreflexives Verhalten zu zeigen (IS und SRIS), als auch die reflektierten Einflussfaktoren (EF2&4) erhöhten sich bei fast allen Teilnehmenden. Das

Verständnis der Begrifflichkeiten *Selbstreflexion* (VS), *Kritisches Denken* (VK) und *kognitive Verzerrungen* (KV) verbesserte sich zudem deutlich und auch die Aspekte der *Epistemologie* (BF-EP) und *Selbstreflexion* (BF-SR) differenzierten sich erkennbar aus.

5.3.3 Interkoderreliabilität

Zur Annäherung an eine Intersubjektivität wurden mindestens 50 % der Datensätze zusätzlich von einem Interkoder mit Expertise in der Chemiedidaktik kodiert und die Interkoderreliabilität ermittelt (s. dazu auch Kap. 4.3.1). Dazu wurden eine eigene unverwendete MAXQDA-Datei, welche die Datensätze enthielt, sowie die notwendigen Kodierleitfäden zur Verfügung gestellt.

Fünf der zehn Datensätze der Interviewstudie S3.I1&2, bestehend aus Pre- und Post-Tests, wurden randomisiert ausgewählt und dem Interkoder zur Verfügung gestellt. Die Kategorien zeigten sehr gute Übereinstimmungswerte ($\varkappa$ = 0,800), vor allem in den durch die vorangegangenen Studien ausgeschärften Kategorien, die sich auf die Aspekte der Bewertung und das Verständnis der Begrifflichkeiten bezogen (Tabelle 100). Die Kategorien der *fachlichen Bewertung* (NW-FB) und der *Fähigkeit zur Bewertung* (FB) erreichten niedrigere, aber dennoch gute Übereinstimmungswerte, was sich auch auf die wenigen sowohl vom Koder als auch Interkoder vergebenen Kodierungen zurückführen lassen könnte.

Tabelle 100: S3.I1&2 – Interkoderreliabilität der Interviewerhebung

Kategorie	Cohens-Kappa ($\varkappa$)
Aspekte naturwissenschaftlicher Bewertung (NW-A)	$\varkappa$ = 0,889
Aspekte der Selbstreflexion (SR-A)	$\varkappa$ = 0,897
Aspekte epistemologischen Verständnisses (EP-A)	$\varkappa$ = 0,880
Verständnis Selbstreflexion (SR-V)	$\varkappa$ = 0,894
Verständnis Kritisches Denken (KD-V)	$\varkappa$ = 0,879
Selbsteinschätzung Fachwissen (NW-SF)	$\varkappa$ = 0,843
Fachliche Bewertung (NW-FB)	$\varkappa$ = 0,724
Fähigkeit zur Bewertung (FB)	$\varkappa$ = 0,720

Ein Beispiel für Stellen, die vom Interkoder und Koder unterschiedlich bewertet wurden, ist das folgende Zitat: „Ja, Menschrechte und sowas, Rassismus, da bin ich sehr emotional involviert. Da bin ich dann nicht ganz so […] neutral“ (S2.I2-SS23-A-B2_3-BN10LB, Pos. 37). Dieses wurde vom Koder in die Kategorie *Emotionen*

(SR-A-IF-EM) der *selbstreflexiven Aspekte* (SR-A) kodiert, da die emotionale Involviertheit explizit hervorgehoben wurde. Zusätzlich dazu ordnete der Interkoder diesem Zitat jedoch die Kategorie *Moral und Werte* (SR-A-IF-MW) zu, da die emotionale Befangenheit auf das Wertesystem der Person zurückzuführen war. In der späteren Kodierkonferenz konnte diese Stelle diskutiert und sich derart geeinigt werden, dass der Vorschlag des Interkoders angenommen wurde.

Ein anderes Beispiel ist das Zitat

> Also, ja einfach wirtschaftliche Aspekte generell. Also wenn sich jetzt eine davon aufgrund von Absätzen deutlich mehr lohnt, dann könnte man vielleicht auch abwägen zu sagen okay, vielleicht ist sie nicht ganz so effizient oder vielleicht ein bisschen weniger umweltfreundlich, aber dadurch, dass sie sich so krass mehr lohnt, wird das so viel mehr ausgebaut (S3.I2-SS23-A-B3_5_DE17RG, Pos. 11).

Dieses wurde vom Koder als *Kostenabwägung* (NW-A-KA) kodiert, da wirtschaftliche Aspekte bei dieser Aussage explizit im Vordergrund standen. Der Interkoder erwog zudem, auch die Kategorien *Nachhaltigkeit* (NW-A-NA) und *Effizienz* (NW-A-EF) zu kodieren, da beide Aspekte im Abwägeprozess benannt werden. In der späteren Kodierkonferenz wurde sich jedoch gegen diese Mehrfachkodierung entschieden, da diese Aspekte in der Aussage nicht im Vordergrund standen, sondern als Beispiel für die Kosten-Nutzen-Abwägung dienten.

Von der Fragebogenerhebung (S3.F1&2) wurde aufgrund des geringeren Datenumfangs im Vergleich zur Interviewerhebung das gesamte Datenmaterial interkodiert. Auch hierbei ergaben sich hohe Übereinstimmungen zwischen Koder und Interkoder, was vermutlich auf die große Übereinstimmung mit den mehrfach überarbeiteten Kodierleitfäden der Interviewstudien zurückzuführen ist (Tabelle 101).

Tabelle 101: S3.F1&2 – Interkoderreliabilität der Fragebogenerhebung

Kategorie	Cohens-Kappa ($\varkappa$)
Bewertungsfaktoren (BF)	$\varkappa = 0{,}870$
Verständnis Selbstreflexion (VS)	$\varkappa = 0{,}894$
Verständnis Kritisches Denken (VK)	$\varkappa = 0{,}879$
Verständnis kognitive Verzerrungen (KV)	$\varkappa = 0{,}843$

Auch bei der Kodierung der offenen Fragen der Fragebögen gab es unterschiedliche Interpretationen seitens des Koders und Interkoders. So wurde die Aussage „Nicht zu allem gleich ja oder nein zu sagen und nicht direkt zuzustimmen, sondern sich erstmal selber mit etwas auseinanderzusetzen, bevor man sich eine Meinung bildet“ (S3.F1-SS23-B3_5_ IP17SE, Pos. 3) vom Koder als *Skepsis* (VK-SK) interpretiert, da eine grundlegend zweifelnde Haltung impliziert wurde. Der Interkoder kodierte hingegen die Kategorie *selbstständiges Denken* (VK-SD). Obwohl beide Kodierungen je nach Interpretation grundsätzlich möglich sind, wurde sich in der Kodierkonferenz für das Beibehalten der Kodierung *Skepsis* (VK-SK) entschieden, da dieser Aspekt zuerst angeführt und davon ausgehend erweitert wurde.

Derlei verschiedene Interpretationsmöglichkeiten verdeutlichen die Herausforderung, die mit der Kodierung solch komplexer Konstrukte einhergeht. Dennoch zeigen die ermittelten Übereinstimmungen eine sehr gute Trennschärfe der einzelnen Kategorien, weshalb diese für weitere Erhebungen als geeignet angesehen werden können.

5.3.4 Rückmeldungen der Teilnehmenden zur Vermittlungseinheit

Zur Überprüfung der Funktionalität des Instruments wurde, wie in Kapitel 4.2.5 beschrieben, eine Vermittlungseinheit im Rahmen der Seminarstruktur eingesetzt und die Methode der Systemisch-reflexiven Stellungnahme (SrS) erprobt. In diesem Rahmen wurde von den Studierenden in der vierten und letzten Aufgabe der SrS eine Prozessreflexion durchgeführt, welche einen Einblick in die Gedanken der Studierenden bezüglich der empfundenen Herausforderungen und Veränderungen während des Prozesses ermöglichen sollte. Folgend soll ein exemplarischer Einblick in die Rückmeldungen der Studierenden gegeben werden[14]. Aufgrund des Umstands, dass die Abgaben zur Reflexionsaufgabe unter Klarnamen erfolgten, konnte keine Zuordnung zu den in den Fragebögen und Interviews verwendeten Kürzeln vorgenommen werden. Es wurden daher vereinfachte numerische Kürzel verwendet.

Die Bedeutung der Selbstreflexion und des Verständnisses kognitiver Prozesse in Bewertungsprozessen wurde von den Teilnehmenden auf verschiedene Weise beschrieben. Dabei traten emotionale, soziale und epistemologische Einflüsse als zentrale Herausforderungen hervor. Ein Aspekt, der genannt wurde, ist die Schwierigkeit,

[14] Die Zitate wurden im Original belassen und auf die Kenntlichmachung von sprachlichen, orthografischen oder grammatischen Fehlern zugunsten der Lesbarkeit verzichtet.

bereits gefasste Meinungen zu reflektieren und zu revidieren, sowie die Notwendigkeit, ein Bewusstsein dafür zu entwickeln, wie folgendes Zitat veranschaulicht:

> Die größte Herausforderung ist das Zugeben bzw. sich das Eingestehen der veränderten Meinung nach neuen Ergebnissen. Wenn man zu einem Thema schon eine bestimmte Meinung und Positionierung hat, ist es eher schwieriger sich vom Gegenteil überzeugen zu lassen. Gerade die Emotionale Ebene macht es einem schwierig. Aber auch hier: Wenn man sich dessen bewusst ist, ist es schon ein großer Schritt in Richtung Offenheit und einem gewissen Abstand zu der persönlichen Emotionalen Ebene und dadurch kann man diese dann besser reflektieren und sich besser neupositionieren (S3-SrS-1).

Ein weiterer zentraler Punkt war die Rolle der affektiven und sozialen Einflüsse, welche die Wahrnehmung und Interpretation von Informationen lenken können. Ein*e Teilnehmende*r stellte die bewusste Reflexion derlei Einflüsse und die Bedeutung einer bewussten Distanz heraus, um diesen Einflüssen zu begegnen:

> Bei der Bearbeitung des Themas ist mir aufgefallen, dass kognitive Verzerrungen und kulturelle Aspekte eine spürbare Rolle spielen. Unsere individuellen Erfahrungen, Überzeugungen und kulturellen Hintergründe beeinflussen, wie wir Informationen aufnehmen und interpretieren. […] Ebenso ist der Einfluss von Familie und Freunden nicht zu unterschätzen, da ihre Meinungen und Perspektiven unser Denken beeinflussen können. Es ist wichtig, sich dieser Einflüsse bewusst zu sein und eine kritische Betrachtung zu pflegen, um eine ausgewogene und objektive Sichtweise zu gewährleisten (S3-SrS-2).

Neben diesen affektiven Faktoren wurde auch die Unsicherheit in Bezug auf die eigene Entscheidung thematisiert. Der verwendete Impulsleitfaden zur Selbstreflexion half den Studierenden bei der Reflexion der eigenen Entscheidungsfindung und dem Hinterfragen der Bedingungen, unter denen die Entscheidungen getroffen wurden:

> Durch den Impulsleitfaden habe ich bemerkt, dass ich mich zwar für eine Meinung entschieden habe, dennoch im Moment noch sehr unsicher mit meiner Entscheidung bin. Ob diese wirklich richtig ist und wenn ich andere Literatur in den 30 Minuten angeschaut hätte, wohl zu einer anderen Entscheidung gekommen wäre (S3-SrS-3).

Auch die mit Socioscientific Issues verbundene Uneindeutigkeit von Bewertungssituationen wurde als Herausforderung benannt, wie das nachfolgende Zitat zeigt:

> Problematisch für die Positionierung – und das ist mir nicht nur bei diesem Thema aufgefallen – ist, dass eine eindeutige und endgültige Aussage eigentlich nicht zu treffen ist. Ganz im Sinne von Natur of Science ist meine Positionierung immer nur vorläufig. Einerseits, weil sich Sachlagen mit neuen Erkenntnissen immer ändern können, anderseits, weil ich nicht einschätzen kann, wie viel ich zu einem Thema tatsächlich weiß (S3-SrS-4).

Darüber hinaus wurde auch die Bedeutung sozialer Normen und externer Erwartungen hervorgehoben. Eine Person reflektierte, wie soziale Faktoren und das Bestreben, der Mehrheit zu entsprechen, die eigene Positionierung erschweren können. Das mit der SrS verbundene dialogische Selbstreflexionstraining schien einen positiven Einfluss darauf zu haben, eine positive Distanz zu diesen externen Faktoren zu wahren, wie folgendes Zitat zeigt:

> Im Hinblick auf meine Selbstreflexion habe ich selbst gemerkt, dass man oft zu bequem handelt und in der Masse „mitschwimmt", obwohl man es selbst eigentlich auch nicht gut findet. […]. In der Erarbeitung des dialogischen Selbstreflexionstrainings ist mir aufgefallen, dass oft Unsicherheiten auftreten können, wie man sich selbst „richtig" positionieren kann, um nicht negativ aufzufallen […]. Insbesondere die Verzerrungen durch soziale Faktoren waren für mich besonders spürbar, weil mich die soziale Norm […] geprägt und mein Weltbild somit auch automatisch mitgestaltet [hat]. […]. Das dialogische Selbstreflexionstraining hat mir hier besonders geholfen, weil ich durch die Strategie, mich auf mich und meine eigenen Werte und Ansichten zu konzentrieren, viel kritischer, selbstsicherer und nachdenklicher geworden bin, im Leben aufzufallen, Stellung zu beziehen, die Hand zu erheben und nicht einfach im Strom „mitzuschwimmen" (S3-SrS-5).

Auch ein*e weitere*r Teilnehmende*r benannte die Relevanz des Erkennens und Bewusstwerdens kognitiver Verzerrungen, um mit diesen umzugehen und diese aktiv in den Entscheidungsprozessen zu berücksichtigen:

> Es ist wichtig, diese Beeinflussungen als kognitive Verzerrungen zu erkennen und sich bewusst zu machen, dass sie meine Entscheidungen beeinflussen können. Der erste Schritt bestand darin, diese Verzerrungen erst einmal überhaupt zu erkennen. Ich habe mich gefragt, ob meine Meinung wirklich auf meinen eigenen Überzeugungen basiert oder ob sie von äußeren Einflüssen geprägt ist. […] Indem ich diese Schritte befolgt habe - das Erkennen, Hinterfragen und Reflektieren - konnte ich besser mit den kognitiven Verzerrungen umgehen, die meinen Entscheidungsfindungsprozess beeinflussten. Ich konnte lernen, meine eigenen Überzeugungen von den Einflüssen anderer zu unterscheiden und fundierte Entscheidungen zu treffen, die meinen eigenen Werten und Normen entsprechen (Sr-SrS-6).

Die Darstellung einer teilnehmenden Person verdeutlicht zudem, dass unbewusste Voreingenommenheit Rechercheprozesse beeinträchtigen und eine nachfolgende Reflexion einen Beitrag zu einer fundierten Positionierung leisten kann:

> Nach der Vorstellung und Besprechung unserer Ergebnisse in der Kleingruppe ist mir aufgefallen, dass meine Positionierung anhand meiner Voreinstellung sehr vorhersehbar war. Ich hatte mir zu Beginn meiner Recherche fest vorgenommen, möglichst objektiv vorzugehen und den Argumenten der Position, zu der ich intuitiv nicht tendiere bzw. gegen die ich einen inneren Widerstand spüre sorgfältig nachzugehen und diese möglichst objektiv zu prüfen. Im Nachhinein glaube ich, dass ich dabei nicht sehr sorgfältig vorgegangen bin, mich

> z.B. nicht weitergehend mit der wirtschaftlichen Lage Deutschlands und den Folgen einer Importabhängigkeit für Erdgas von Amerika Expertise befasst habe […] (S3-SrS-7).

Auch in den Interviews gab es Aussagen, die sich auf die Vermittlungseinheit und die Methode der SrS bezogen. Die nachfolgende Äußerung knüpft thematisch am vorangegangenen Zitat an und spiegelt den Einfluss der subjektiv empfundenen Objektivität wider:

> Weil ich zu Beginn an eigentlich dachte, dass ich mega kritisch vorgehe und jetzt auf jeden Fall versuche, mir eine objektive Meinung zu bilden und möglichst alles in Betracht zu ziehen, aber ich bei dieser Reflexionsaufgabe dann gemerkt habe, dass [ich] eigentlich gar nicht so kritisch vorgegangen bin und es so sehr vorausschaubar war, wie ich mich am Ende positioniere (S3.I2-SS23-B-B3_5_SJ16DI-Post, Pos. 45).

Diese Perspektiverweiterung wird auch bei einer anderen teilnehmenden Person deutlich:

> [I]ch bin mir relativ sicher, dass ich beim letzten Mal auch gesagt habe, dass ich ein sehr reflektierter Mensch bin. Und dann im Rahmen des Seminars ist mir halt eigentlich aufgefallen, dass man halt sogar noch tiefer in die Ebene reingehen konnte des Reflektierens, und ich mir dann so zeitweise dachte, ach krass, okay, so, das geht ja hier noch viel tiefer. Und war ich davor überhaupt so krass reflektiert, wie ich jetzt dachte? Oder dachte ich das einfach nur in dem Moment, weil ich es nicht besser wusste (S3.I2-SS23-B-B3_5_XJ36JE, Pos. 51)?

Den empfundenen Einfluss der Seminarinhalte benannten auch andere Teilnehmende:

> Ich würde sagen, vor dem Seminar hätte ich das fast gar nicht gemacht. Und seitdem ich das Seminar besucht habe, würde ich da viel mehr drauf achten (S3.I2-SS23-A-B3_5_XF21FM, Pos. 29).

> Ich habe aber schon das Gefühl, dass jetzt, im Vergleich zum ersten Interview, also zu dem Zeitpunkt ich noch einmal viel ein klareres Bild habe, wie man mit Informationen umgeht und mir nochmal viel bewusster bin über die ganzen Verzerrungen, die es geben kann. Und es auch, also ich auch automatischer daran denken (S3.I2-SS23-B-B3_5_SJ16DI, Pos. 47).

> [D]urch das Seminar haben wir noch mal, da bin ich ehrlich: Kultur war immer so, ja man hat das Wort gehört und man weiß ungefähr, was das ist, aber was wirklich Kultur für einen Einfluss hat, hatte ich vorher keine Ahnung von. Und jetzt habe ich das noch mal so mehr gemerkt, weil wir halt auch darüber gesprochen haben. [F]ür uns sind Dinge normal oder das sollte so sein und in anderen Kulturen halt gar nicht. Und dann habe ich noch mal gemerkt, wie unterschiedlich das ist, wo man halt auch aufgewachsen ist und welche Normen und Werte man bekommen hat (S3.I2-SS23-B-B2_3-SV32BT, Pos. 43).

> [W]ir haben ja auch nochmal darüber geredet in dem Seminar, dass man auch vielleicht noch mal bewusstwird, welche Quellen jetzt noch mehr, mehr seriöser sind als andere. Das fand ich auch noch mal ganz gut, weil man sich dann selber noch mal ins Gehirn ruft und noch mal selber überdenkt (S3.I2-SS23-B-B3_5_IP17SE, Pos. 41).

> Also mir ist auf jeden Fall bewusster geworden, gerade so Internetquellen noch mal wirklich genau zu hinterfragen und auch Zeitschriften, also Onlinezeitschriften, genauer zu hinterfragen. Aber auch diese kognitiven Verzerrungen waren mir vorher gar nicht so bewusst. Und jetzt nehme ich die viel öfter wahr (S3.I2-SS23-A-B2_3-XJ189MF, Pos. 21).

Die Reflexion von Einflüssen scheint besonders bei der Auseinandersetzung mit rein naturwissenschaftlichen und chemischen Themen eine besondere Rolle zu spielen, da ein Lernen fachlicher Inhalte nicht per se auch zu einer Anwendung des Wissens in der Lebensrealität zu führen scheint, wie folgende Äußerung nahelegt:

> [A]lso ich finde jetzt auch, eben auch so ein naturwissenschaftlicher Aspekt oder so was, jetzt auch im Chemiestudium oder so, so viel wirklich davon nehme ich jetzt auch jetzt nicht mit für solche Themen. Also, das ist ja viel so, was wir in der Chemie wir machen, ja auch sehr klein. Das ist ja jetzt nicht, das trifft ja jetzt keine großen gesellschaftlichen Themen irgendwie (S3.I2-SS23-B-B3_5_IP17SE, Pos. 39).

Gerade das Erkennen von naturwissenschaftlichen Themen in der Lebensrealität ist aber ein essenzieller Teil naturwissenschaftlicher Grundbildung und bedarf daher neben dem Wissen der Fähigkeit zur Reflexion:

> Desto mehr reflektiere ich mich auch, weil wenn ich meine Meinung sage, ich das halt auch mit gutem Gewissen und fundiert sagen möchte. Und dafür brauche ich halt Reflexion (S3.I3-SS23-A-B2_3-NB10LB, Pos. 42).

Abschließend soll hier noch ein Zitat angeführt werden, welches die Relevanz selbstreflexiver Denkweisen besonders in Verbindung mit den Naturwissenschaften hervorhebt:

> Ich würde sagen, es ist auf jeden Fall eine Fähigkeit, mit sich mittlerweile auch viel reflektierender und auch mit seinem Wissen umzugehen, zu einer Entscheidung zu kommen und diese zu lösen. In sich reinzuhorchen und sich zu fragen, wie denke ich eigentlich darüber? Ich wusste zu wenig darüber und ich habe mich zu wenig damit befasst und auch zu wenig darüber gelernt. Man hat immer gedacht, oh ja, man studiert Chemie und man kann kritisch denken irgendwie, aber eigentlich kann man das dann doch nicht. Und das ist einem, finde ich auch, gerade dann erst mal bewusst geworden. Man beschäftigt sich viel mehr mit sich selbst, finde ich jetzt. Also ich finde, man ist irgendwie, viel mehr über sich hinausgewachsen (EH2-SS23-A-B3_5_XF21FM-Post, Pos. 51).

Entgegen der Vorstellung einer durch das Lernen von Fachwissen implizit erworbenen Fähigkeit zum Kritischen Denken unterstreicht diese Aussage die Wichtigkeit der Vermittlung eines multiperspektivischen Denkansatzes, welcher selbstreflexive Denkweisen bewusst inkludiert, um das naturwissenschaftliche Wissen zu ergänzen und damit zu stärken.

5.3.5 Abschluss der Studie S3

Die Resultate der Studie S3 weisen auf eine deutliche Veränderbarkeit der Denkprozesse in Bewertungssituationen durch eine explizite Thematisierung Kritischen Denkens hin, welche mithilfe des entwickelten Instrumentariums dargestellt werden konnte. Daraus lassen sich zwei mögliche Schlussfolgerungen ableiten:

Zum einen scheint die zugrundeliegende konzeptionelle Entwicklung das Instrument derart optimiert zu haben, dass die in den Seminareinheiten angestrebte Förderung Kritischen Denkens auch durch dieses operationalisiert und mittels des Vergleichs zwischen Pre- und Post-Tests gemessen werden konnte. Zum anderen legt dies die Annahme nahe, dass Kritisches Denken durch explizite Thematisierung grundsätzlich adressierbar zu sein scheint.

Die Ergebnisse zeigen dabei deutliche Verbesserungen von sowohl Fremd- als auch Selbsteinschätzung, was zum einen für eine verbesserte Selbstwahrnehmung und subjektiv empfundene Steigerung der Selbstreflexionsfähigkeit spricht, zum anderen aber auch tatsächliche Perspektiverweiterungen in konkreten Bewertungsprozessen widerspiegelt. Es kann demnach nicht nur eine Steigerung des individuellen Verständnisses auf Ebene des Faktenwissens, sondern durch die Fremdeinschätzung auch eine Kompetenzerweiterung auf Ebene des Verfügungswissens verzeichnet werden.

Diese Ergebnisse sind jedoch mit Vorbehalt zu interpretieren, da der beobachtete Effekt, insbesondere im Hinblick auf das Verständnis der zugrundeliegenden Konstrukte, aufgrund der expliziten Thematisierung innerhalb der Seminareinheiten zu erwarten war. Gleichzeitig sind das Wissen und Verständnis eines Begriffs nicht gleichbedeutend mit der Anwendung des Gelernten, auch wenn das Verstehen einen Beitrag zur Umsetzung in Handeln leisten kann. In der Folge bleibt eine verknüpfte Betrachtung von Fremd- und Selbsteinschätzung notwendig, um beide Ebenen in Relation zueinander bewerten zu können.

Dennoch zeigen besonders die Ergebnisse der Interviewstudie Anzeichen für nachhaltige Verbesserungen der Bewertungsperspektiven in konkreten Kontexten mit naturwissenschaftlichem Bezug und damit auch Ausprägungen auf Kompetenzebene. Vor allem die Resultate der Follow-Up-Interviews deuten darauf hin, dass die Teilnehmenden nicht nur kurzfristig von der expliziten Thematisierung Kritischen Denkens profitieren konnten, sondern dass die gewonnenen Erkenntnisse und Strategien auch langfristig in alltägliche Bewertungs- und Entscheidungsprozesse integriert werden konnten. Diese Annahme wird zudem von den Studierendenfeedbacks unterstützt, welche ebenfalls erkennen lassen, dass die bewusste Auseinandersetzung mit Kritischem Denken, insbesondere in Verbindung mit konkreten Bewertungsanlässen und Herangehensweisen, auch auf individueller Ebene positive Impulse setzen konnte.

Das Instrumentarium erweist sich daher in seiner kombinierten Form von Fremd- und Selbsteinschätzung als geeignet, um sich diesem komplexen Konstrukt nähern zu können, weshalb dessen Struktur für weitere Studien beibehalten wurde.

Genauer betrachtet zeigte sich die Interviewform besonders für die Fremdeinschätzung als zielführend, während die konkreten Bewertungsanlässe eine fundierte Analyse der Bewertungsprozesse ermöglichten. Die Ergebnisse der unterschiedlichen eingesetzten Settings zeigten darüber hinaus keine feststellbaren Unterschiede, was eine Kontextunabhängigkeit nahelegt. Eine Überprüfung dessen fand in nachfolgenden Studien gezielt statt.

Ungerichtete Erhebungsansätze in Form der offenen Fragen zu den *Bewertungsfaktoren* (BF) des Fragebogens, welche keine konkreten Settings beinhalteten, erwiesen sich als erkennbar weniger ergiebig. Aus diesem Grund kann angenommen werden, dass eine Erweiterung der offenen Fragestellungen des Fragebogens um konkrete Bewertungssituationen eine sinnvolle Ergänzung zur Erhebung einer breiteren Stichprobengröße bietet. Die Erprobung dieser Ergänzung wurde in Studie S4 umgesetzt.

Zur Überprüfung der Reproduzierbarkeit der Ergebnisse und Sicherstellung der Kontextunabhängigkeit des Instruments wurde eine Wiederholung dieses Erhebungsaufbaus mit einer weiteren Stichprobe angestrebt.

5.4 Studie S4

Die vierte Studie S4 diente als Zwischenstudie zur Überprüfung der Kontextunabhängigkeit und Stabilität des in Studie S3.F1&2 eingesetzten Fragebogens. Die Datenerhebung fand am Ende des Sommersemesters 2023 im Rahmen des Blockseminars zur *Einführung in das wissenschaftliche Arbeiten* statt, das im Basismodul *Grundlegende Aspekte der Fachdidaktik (B2 HR/SP B3 GG)* verankert ist und vom Untersuchenden in seiner Funktion als Dozent abgehalten wurde. Insgesamt nahmen zwei Gruppen an den Blockveranstaltungen teil, was eine vergleichende Erhebung zweier unabhängiger Gruppen ermöglichte. In der Gruppe S4.F1 wurden $N = 17$ vollständige Datensätze gewonnen, während in der Gruppe S4.F2 $N = 14$ vollständige Datensätze erhoben wurden (Abbildung 25).

Abbildung 25: Übersicht über den Aufbau und die Erhebungsparameter der Studie S4

Der Aufbau der eingesetzten Fragebögen entsprach dem in Studie S3.F1&2 eingesetzten Erhebungsinstrument, welches basierend auf den in dieser Studie gewonnenen Ergebnissen um kontextualisierte Fragestellungen zu relevanten Bewertungsaspekten erweitert wurde. Dazu wurden vier kontextualisierte offene Fragen nach relevanten *Bewertungsaspekten* (ASP) hinzugefügt, welche die bisherigen nicht-kontextualisierten offenen Fragestellungen zu den *Einflussfaktoren* (EF) ergänzten.

Zur Kontextualisierung wurden für den Fragenblock *Bewertungsaspekte* (ASP) Settings vordefiniert, von denen ausgehend die Beantwortung der Fragen erfolgen sollte (Tabelle 102). Diese Settings variierten zwischen den Gruppen S4.F1 und S4.F2, um die Untersuchung des Einflusses der Kontexte auf die Bewertung und etwaige Gruppenunterschiede zu ermöglichen. Für die Fragebogenerhebung in Studie S4.F1&2 wurden die bereits eingesetzten Kontexte Solarenergieversorgung vs. Windkraftanlage (Kontext A) und E-Mobilität vs. Brennstoffzellenantrieb (Kontext B) verwendet (vgl. Kap. 4.2.6.2 und 4.2.6.3).

Tabelle 102: S4.F1&2 – Settings des Blocks *Bewertungsaspekte* (ASP)

Gruppe	Setting
Setting S4.F1	Folgendes Setting definiert die Ausgangslage: Sie werden für eine politische Umfrage befragt, die für die zukünftige Mobilitätsausrichtung des Landes relevant ist. Aus Klimaschutzgründen wird der Verbrenner abgeschafft und es stehen zwei alternative Antriebsmöglichkeiten zur Diskussion: die Elektromobilität und der Wasserstoffantrieb. Die Frage, die nun im Raum steht, ist folgende: – Sollte der Ausbau und die Förderung von Elektrofahrzeugen verstärkt werden oder – sollte der Ausbau und die Förderung von Brennstoffzellenautos verstärkt werden, um den Einsatz von fossilen Brennstoffen im Verkehrssektor zu reduzieren?
Setting S4.F2	Folgendes Setting definiert die Ausgangslage: Sie werden für eine politische Umfrage befragt, die für die zukünftige Energiepolitik des Landes relevant ist. Aus Klimaschutzgründen werden Energiegewinnungsanlagen, welche fossile Brennstoffe benötigen, abgeschafft und es stehen zwei alternative Energiegewinnungsmöglichkeiten zur Diskussion: die Energiegewinnung mittels Windkraftanlagen oder durch Solarenergie. Die Frage, die nun im Raum steht, ist folgende: – Sollte der Ausbau von Windenergieanlagen vorrangig vorangetrieben werden oder – sollte der Ausbau von Solarenergieanlagen verstärkt gefördert werden, um den Einsatz von fossilen Brennstoffen im Energiesektor zu reduzieren?

Die Fragestellungen der *Bewertungsaspekte* (ASP) orientierten sich an den Fragestellungen der Interviewleitfäden, wurden jedoch an das Erhebungsformat der Fragebögen angepasst und daher gezielter formuliert (Tabelle 103).

Tabelle 103: S4.F1&2 – Fragestellungen des Blocks *Bewertungsaspekte* (ASP)

Fragencode	Fragestellung
ASP1	Wie würden Sie sich entscheiden und warum? Bitte führen Sie alle für Sie relevanten Aspekte an, um sich klar positionieren zu können. (Sie müssen nicht bereits alle Informationen, die für Sie relevant wären, vorliegen haben. In diesem Falle stellen Sie gerne Fragen nach den für Sie relevanten Aspekten.)
ASP2	Welche Aspekte wären aus fachlicher bzw. naturwissenschaftlicher Sicht für Sie relevant, um sich klar positionieren zu können?
ASP3	Welche Aspekte wären mit Blick auf sich selbst für Sie relevant, um sich klar positionieren zu können?
ASP4	Welche Aspekte sind für Sie relevant, wenn Sie Quellen hinsichtlich ihrer Glaubwürdigkeit einschätzen müssen?

Die in ASP1 formulierte Fragestellung erfasste, ähnlich wie die Interviewleitfäden in Studie S3.I1&2, die als relevant empfundenen Aspekte zunächst ungerichtet, bevor in ASP2 bis 4 gezielt die einzelnen Teilbereiche adressiert wurden.

Der Aufbau des vollständigen Fragebogens, einschließlich des neu eingefügten Fragenblocks zu den *Bewertungsaspekten* (ASP), ist in Tabelle 104 dargestellt.

Tabelle 104: S4.F1&2 - Abfolge der Frageblöcke im finalen Fragebogen

Fragencode	Fragestellung	Itemformat
PI	Persönliche Informationen	
ASP	Bewertungsaspekte	Offen
VF1	Verständnisfrage 1	Offen
VF2	Verständnisfrage 2	Offen
CRT1	Erste randomisierte Frage des CRT	Zahleneingabe
VF3	Verständnisfrage 3	Offen
IS	Introspektion	Geschlossen
SR	SRIS – Self-Reflection	Geschlossen
CRT2	Zweite randomisierte Frage des CRT	Zahleneingabe
IN	SRIS - Insight	Geschlossen
EF1	Einflussfaktoren 1	Offen
CRT3	Dritte randomisierte Frage des CRT	Zahleneingabe
EF2-4	Einflussfaktoren 2-4	Offen/Geschlossen
BG	Bewertung und Gewichtung	Geschlossen

Die finale Form des Fragebogens umfasste insgesamt neun offene Fragestellungen, 46 geschlossene Likert-Items sowie die drei Aufgaben des CRT und kann im Anhang (A.1.2) eingesehen werden. Die Erweiterung um vier zusätzliche kontextualisierte offene Fragestellungen zu den *Bewertungsaspekten* (ASP) beeinträchtigte die Bearbeitungszeit nicht erheblich, sodass diese weiterhin innerhalb eines zumutbaren Rahmens blieb.

5.4.1 Auswertung

Die Auswertung der Items des Fragebogens erfolgte in Anlehnung an die Fragebogenerhebung S3.F1&2. Eine detaillierte Darstellung der Auswertungsmethodik ist in Kapitel 5.2.2.1 zu finden. Zudem wurden die neuen kontextualisierten offenen Fragestellungen mit dem bereits eingesetzten Kodierleitfaden zu den *Bewertungsfaktoren*

(BF) kodiert, welcher ebenfalls in Kapitel 5.2.2.1 erläutert und im Detail im Anhang (A.3.1.2) dokumentiert ist.

Zusätzlich zur Auswertung der Antworten der Fragebogendaten wurden die qualitativen Datensätze der offenen Fragestellungen erneut von einem Interkoder mit Expertise auf dem Gebiet der Chemiedidaktik interkodiert und die Interkoderreliabilität errechnet. Ferner wurden die beiden Gruppen S4.F1 und S4.F2 mittels des Mann-Whitney-U-Tests verglichen, um mögliche Unterschiede in den Gruppenvoraussetzungen sowie potenzielle Effekte durch die unterschiedlichen verwendeten Kontexte zu analysieren.

5.4.2 Ergebnisse

Da die Studie S4 als Zwischenstudie konzipiert war und primär der Überprüfung der Stabilität und Kontextunabhängigkeit des Instruments diente, wird auf eine detaillierte Darstellung der einzelnen Gruppenergebnisse verzichtet. Stattdessen werden ausgewählte Ergebnisse vorgestellt, die von besonderer Relevanz für die Optimierung des Instruments waren.

Die implementierte Kontextualisierung der offenen Fragen zu den *Bewertungsaspekten* (ASP) zeigte im Vergleich zu Studie S3.F1&2 ein deutlich verbessertes Antwortverhalten im Hinblick auf die benannten Bewertungsperspektiven, insbesondere auf *epistemologischer* (BF-EP) und *naturwissenschaftlicher* (BF-NW) Ebene.

Im Vergleich zum geringeren Antwortumfang hinsichtlich der *epistemologischen Faktoren* (BF-EP) in der Fragebogenstudie S3.F1&2 zeigten beide Gruppen in der Studie S4.F1&2 eine größere Anzahl an Antworten mit epistemologischem Bezug (Tabelle 105). Dabei wurde deutlich, dass für beide Gruppen die *Herkunft der Informationen* (BF-EP-HK) sowie die *Expertise der Informationsquelle* (BF-EP-EX) eine vornehmliche Rolle einnahmen. Auch die *Transparenz* (BF-EP-TR) im Umgang mit Quellen spielte in beiden Gruppen eine Rolle, zeigte sich jedoch bei Gruppe S4.F1 stärker.

Tabelle 105: S4.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *epistemologische Faktoren* (EP-BF), *N* = 17 und *N* = 14

			S4.F1		S4.F2	
Kat.	**1.UK 2. UK**	**3. UK**	**Summe Einzel**	**Summe Mehrf.**	**Summe Einzel**	**Summe Mehrf.**
BF: Bewertungfaktoren	**EP: Epistemologische Faktoren**	BF-EP-EX: Autoritäten/Expertise	7	7	6	7
		BF-EP-AK: Aktualität	6	7	1	1
		BF-EP-IK: Interessenskonflikte	5	5	2	2
		BF-EP-PR: Präsentation	3	3	8	12
		BF-EP-HK: Herkunft der Information	8	8	10	13
		BF-EP-ME: Methodik	4	5	0	0
		BF-EP-KO: Konsistenzprüfung	3	3	4	4
		BF-EP-TR: Transparenz	7	7	4	4
		BF-EF-SE: Soziokulturelle Eingebundenheit	0	0	0	0
		Summe EP	**43**	**45**	**35**	**43**
		Einzelzählung EP	**8**	**8**	**7**	**7**

Die *Aspekte naturwissenschaftlicher Bewertung* (BF-NW), welche in Studie S3.F1&2 noch sehr schwach ausgeprägt waren, traten nach der Implementierung der Kontexte deutlich stärker hervor (Tabelle 106). Beide Gruppen legten dabei den Schwerpunkt unter anderem auf die *Umsetzbarkeit* (BF-NW-UM) und die *Effizienz* (BF-NW-EF) der Technologien. Auch die *Abwägung der Kosten* (BF-NW-KA) war bei beiden Gruppen ein relevantes Bewertungskriterium. Unterschiede in den Ausprägungen waren bei den Kategorien *Auswirkungen und Nachhaltigkeit* (BF-NW-NA) sowie der *Rohstoffherkunft* (BF-NW-RH) zu verzeichnen, welche zwar in beiden Gruppen benannt wurden, aber unterschiedlich ausgeprägt auftraten. Dies könnte auf unterschiedliche Gewichtungen einzelner Bewertungsaspekte zurückzuführen sein, die durch die jeweiligen Kontexte beeinflusst werden.

Tabelle 106: S4.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *naturwissenschaftliche Faktoren* (BF-NW), *N* = 17 und *N* = 14

			S4.F1		S4.F2	
Kat.	**1.UK 2. UK**	**3. UK**	**Summe Einzel**	**Summe Mehrf.**	**Summe Einzel**	**Summe Mehrf.**
BF: Bewertungsfaktoren	**NW: Naturwissenschaftliche Faktoren**	BF-NW-WS: Wissenschaftliche Standards	2	3	1	1
		BF-NW-EF: Effizenz	9	16	9	13
		BF-NW-KA: Kostenabwägung	11	13	11	17
		BF-NW-NA: Auswirkungen & Nachhaltigkeit	14	32	5	10
		BF-NW-FT: Funktionsweise	3	4	4	6
		BF-NW-UM: Umsetzbarkeit	12	22	12	22
		BF-NW-RH: Rohstoffherkunft	9	16	3	4
		BF-NW-NO: Notwendigkeit	0	0	0	0
	Summe NW		**60**	**106**	**45**	**73**
	Einzelzählung NW		**7**	**7**	**7**	**7**

Durch die konkreten Anforderungssituationen konnte somit ein umfassenderes Bild hinsichtlich naturwissenschaftlicher und epistemologischer Aspekte gewonnen werden. Diese Ergebnisse deuten darauf hin, dass eine explizite Kontextualisierung der Fragestellung notwendig ist, um die Bewertungsperspektiven adäquat zu untersuchen.

Die Überprüfung der Antworten der offenen Fragen auf Gruppenunterschiede mittels des Mann-Whitney-U-Tests zeigte keine signifikanten Differenzen zwischen der Gruppe S4.F1, die den Kontext E-Mobilität vs. Brennstoffzellenantrieb behandelte, und der Gruppe S4.F2, die sich mit dem Kontext Solarenergieversorgung vs. Windenergie auseinandersetzte (Tabelle 107). Weder hinsichtlich der im Bewertungsprozess herangezogenen Perspektiven noch im Hinblick auf das in beiden Gruppen vorherrschende Verständnis der Begrifflichkeiten konnten signifikante Unterschiede festgestellt werden.

Tabelle 107: S4.F1&2 - Ergebnisse der Fragebogenerhebung, Mann-Whitney-U-Test der offenen Fragen zum *Verständnis* (VS, VK, KV) und zu den *Bewertungsfaktoren* (BF), $N = 17$ und $N = 14$

S.4.F1&2 - Bewertungsaspekte (BF), Mann-Whitney-U-Test

Teststatistikena

	S4_VS	S4_VK	S4_BF_WI	S4_BF_EP	S4_BF_SR	S4_BF_SR_IF	S4_BF_SR_EF	S4_KV
Mann-Whitney-U-Test	114,500	107,000	104,000	114,500	112,000	110,000	116,500	108,000
Wilcoxon-W	219,500	212,000	209,000	267,500	217,000	263,000	221,500	213,000
Z	-0,196	-0,581	-0,622	-0,184	-0,286	-0,366	-0,111	-0,549
Asymp. Sig. (2-seitig)	0,844	0,561	0,534	0,854	0,775	0,714	0,911	0,583
Exakte Sig. [2*(1-seitige Sig.)]	,860b	,653b	,570b	,860b	,799b	,739b	,922b	,681b
Exakte Sig. (2-seitig)	0,839	0,597	0,556	0,864	0,790	0,728	1,000	0,643
Exakte Sig. (1-seitig)	0,439	0,311	0,280	0,433	0,405	0,363	0,529	0,341
Punkt-Wahrscheinlichkeit	0,017	0,093	0,006	0,009	0,010	0,008	0,146	0,131
a. Gruppenvariable: Studie_Gruppe								
b. Nicht für Bindungen korrigiert.								
Effektstärke	0,025	0,074	0,079	0,023	0,036	0,046	0,014	0,070

Auch die Ergebnisse der geschlossenen Items beider Gruppen offenbarten keine signifikanten Gruppenunterschiede (Tabelle 108). Lediglich im Itemblock der *Einflussfaktoren* (EF4) wiesen drei Items unterschiedliche Ausprägungen auf. Die Reflexion der *eigenen Bildung* (EF4_5) fiel mit einer Teststatistik von $z = -2{,}485$ und einer Signifikanz von $p = 0{,}013$ in beiden Gruppen unterschiedlich aus, erreichte jedoch nur knapp eine mittlere Effektstärke ($r = 0{,}316$). Ein ähnliches Muster zeigte sich bei der Reflexion der *Motivation* (EF4_18), die mit $z = -2{,}005$, $p = 0{,}045$ und $r = 0{,}255$ lediglich eine schwache Effektstärke aufwies. Auch der Einfluss von *Gruppendynamiken* (EF4_21) wurde in den Gruppen unterschiedlich stark reflektiert ($z = -2{,}449$, $p = 0{,}014$) und erreichte mit $r = 0{,}311$ knapp eine mittlere Effektstärke. Abgesehen von diesen drei Items waren keine signifikanten Unterschiede zwischen den beiden Gruppen zu verzeichnen. Da es sich bei diesem Itemblock um nicht-kontextualisierte, geschlossene Fragen handelte, liegt die Annahme nahe, dass die Ergebnisse auf individuelle Schwerpunktsetzungen der Einzelnen innerhalb der Gruppen zurückzuführen sind.

Tabelle 108: S4.F1&2 - Ergebnisse der Fragebogenerhebung, Mann-Whitney-U-Test der geschlossenen Items, $N = 17$ und $N = 14$

S4.F1&2 - Übersicht Mann-Whitney-U-Test				
	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
IS1	0,308	-1,02	0,130	Nicht signifikant
IS2	0,494	-0,68	0,087	Nicht signifikant
IS3	0,966	-0,04	0,005	Nicht signifikant
IS4	0,833	-0,21	0,027	Nicht signifikant
IS5	0,414	-0,82	0,104	Nicht signifikant
EF2	0,338	-0,96	0,122	Nicht signifikant
EF4_1	0,720	-0,36	0,046	Nicht signifikant
EF4_2	0,310	-1,01	0,129	Nicht signifikant
EF4_3	0,518	-0,65	0,082	Nicht signifikant
EF4_4	0,753	-0,32	0,040	Nicht signifikant
EF4_5	0,013	-2,48	0,316	Signifikant
EF4_6	0,227	-1,21	0,154	Nicht signifikant
EF4_7	0,701	-0,38	0,049	Nicht signifikant
EF4_8	0,935	-0,08	0,010	Nicht signifikant
EF4_9	0,379	-0,88	0,112	Nicht signifikant
EF4_10	0,775	-0,29	0,036	Nicht signifikant
EF4_11	0,616	-0,50	0,064	Nicht signifikant
EF4_12	0,151	-1,44	0,182	Nicht signifikant
EF4_13	0,141	-1,47	0,187	Nicht signifikant
EF4_14	0,455	-0,75	0,095	Nicht signifikant
EF4_15	0,647	-0,46	0,058	Nicht signifikant
EF4_16	0,258	-1,13	0,144	Nicht signifikant
EF4_17	0,142	-1,47	0,187	Nicht signifikant
EF4_18	0,045	-2,00	0,255	Signifikant
EF4_19	0,574	-0,56	0,071	Nicht signifikant
EF4_20	0,400	-0,84	0,107	Nicht signifikant
EF4_21	0,014	-2,45	0,311	Signifikant
EF4_22	0,195	-1,29	0,164	Nicht signifikant
EF4_23	0,315	-1,00	0,128	Nicht signifikant
EF4_24	0,753	-0,31	0,040	Nicht signifikant
EF4_25	0,748	-0,32	0,041	Nicht signifikant
EF4_26	0,697	-0,39	0,049	Nicht signifikant
BG1	0,174	-1,36	0,173	Nicht signifikant
BG2	0,688	-0,40	0,051	Nicht signifikant
SR3	0,279	-1,08	0,138	Nicht signifikant
SR5	0,232	-1,20	0,152	Nicht signifikant
SR6	0,383	-0,87	0,111	Nicht signifikant
SR8	0,887	-0,14	0,018	Nicht signifikant
SR9	0,408	-0,83	0,105	Nicht signifikant
SR10	0,355	-0,93	0,117	Nicht signifikant
IN2	0,777	-0,28	0,036	Nicht signifikant
IN4	0,128	-1,52	0,193	Nicht signifikant
IN5	0,607	-0,51	0,065	Nicht signifikant
IN6	0,808	-0,24	0,031	Nicht signifikant
IN7	0,252	-1,15	0,145	Nicht signifikant
IN8	0,185	-1,32	0,168	Nicht signifikant
CRT1	,200b	0,19	0,024	Nicht signifikant
CRT2	0,193	,200b	0,025	Nicht signifikant
CRT3	0,099	0,19	0,025	Nicht signifikant

Die übrigen Ergebnisse des Fragebogens werden im Folgenden in einem kurzen Überblick dargestellt:

Beide Gruppen interpretierten den Begriff der *Selbstreflexion* (VS) primär als die Reflexion *eigener Handlungen* (VS-ET-HA) und *Gedanken* (VS-ET-IS-GE). *Kritisches Denken* (VK) wurde von der Mehrheit beider Gruppen als das *Hinterfragen* (VK-NA-HF) von Dingen verstanden, wobei diese Definition meist ohne weitere Konkretisierung angegeben wurde. *Kognitive Verzerrungen* (KV) wurden überwiegend als *Denk- und Wahrnehmungsfehler* (KV-TG-DF) beschrieben.

In den jeweiligen Bewertungskontexten reflektierten beide Gruppen auf der *selbstreflexiven Ebene* (BF-SR) primär *interne Faktoren* (BF-SR-IF), wie das *eigene Fachwissen* (BF-SR-IF-FW), das persönliche *Interesse* (BF-SR-IF-UN) sowie mögliche *Befangenheiten* (BF-SR-IF-BV). Als einflussreichster *externer Faktor* (BF-SR-EF) wurde das *soziale Umfeld* (BF-SR-EF-SN) wahrgenommen. Im Bereich der *epistemologischen Aspekte* (BF-EP) erwiesen sich die *Herkunft der Informationen* (BF-EP-HK) und die *Expertise* (BF-EP-EX) der Quelle als wichtigste Bewertungsfaktoren. Auch auf *naturwissenschaftlicher Ebene* (BF-NW) waren die benannten relevanten Faktoren trotz unterschiedlicher Kontexte sehr vergleichbar und konzentrierten sich auf die *Umsetzbarkeit* (BF-NW-UM) und die *Abwägung der Kosten* (BF-NW-KA).

Bei den geschlossenen Items zur *Introspektion* (IS) erreichten beide Gruppen über alle Items hinweg vergleichbare Mittelwerte, mit einem Gesamtmittelwert von 3,412 Skalenpunkten für Gruppe S4.F1 und 3,243 für Gruppe S4.F2. Diese Ergebnisse deuten auf eine moderate Reflexionsbereitschaft interner Faktoren, wie Gefühle, Moralvorstellungen oder das Interesse, hin.

Als besonders häufig reflektierte *Einflussfaktoren* (EF4) nannten beide Gruppen vorwiegend das *eigene Wissen* (EF4_12) und den *medialen Kontext* (EF4_13). Die eigene *religiöse Überzeugung* (EF4_9) spielte für die Mehrheit der Teilnehmenden eine untergeordnete Rolle.

Im Itemblock *Bewertung und Gewichtung* (BG) lagen die Mittelwerte der Gruppe S4.F1 bei 2,765 und die der Gruppe S4.F2 bei 2,500. Beide Gruppen bewerteten damit die Einflussfaktoren aus EF4 insgesamt als moderat bis schwach und gaben an, diese nur gelegentlich zu reflektieren.

Der SRIS machte bei Gruppe S4.F1 mit einem Mittelwert von 5,402 eine leicht höhere Tendenz, *selbstreflexives Verhalten* (SR) zu zeigen, sichtbar als bei Gruppe S4.F2

mit einem Mittelwert von 4,964. Beide Gruppen erreichten auf der Skala der *Einsicht* (IN) niedrigere Werte mit durchschnittlich 4,000 bei Gruppe S4.F1 und 4,369 bei Gruppe S4.F2.

Im CRT wurde eine ähnliche Verteilung der Ergebnisse, welche auch mit den Vorstudien vergleichbar waren, erkenntlich (Tabelle 109).

Tabelle 109: S4.F1&2 – Ergebnisse der Fragebogenerhebung, CRT, $N = 17$ und $N = 14$

S4.F1			
Verteilung der Ergebisse in %			
17,6	23,5	35,3	23,5
0	1	2	3
Low			High

S4.F2			
Verteilung der Ergebisse in %			
14,3	28,6	28,6	28,6
0	1	2	3
Low			High

5.4.3 Interkoderreliabilität

Auch in der Studie S4 wurde die Kodierung zusätzlich von einem weiteren Doktoranden der Chemiedidaktik vorgenommen und zur weiteren Überprüfung der Intersubjektivität die Interkoderreliabilität bestimmt (Tabelle 110)(vgl. Kap. 4.3.1). Insgesamt wurden neun der 17 Datensätze (52,9 %) der Gruppe S4.F1 und sieben der 14 Datensätze (50 %) der Gruppe S4.F2 randomisiert ausgewählt und dem Interkoder zur Verfügung gestellt.

Tabelle 110: S4.F1&2 – Interkoderreliabilität der Fragebogenerhebung

Kategorie	Cohens-Kappa ($\varkappa$)
Bewertungsfaktoren (BF)	$\varkappa = 0{,}898$
Verständnis Selbstreflexion (VS)	$\varkappa = 0{,}890$
Verständnis Kritisches Denken (VK)	$\varkappa = 0{,}871$
Verständnis kognitive Verzerrungen (KV)	$\varkappa = 0{,}897$

Die Übereinstimmungen der Kodierungen und die errechnete Interkoderreliabilität zeigten insgesamt vergleichbare Werte zu Studie S3.F1&2. Die Kodierungen aller vier Kodierleitfäden erreichten hohe Übereinstimmungen, was für eine hohe Trennschärfe der entwickelten Kategorien spricht und sich auf die iterative Überarbeitung

dieser innerhalb der ersten Studien zurückführen lässt. Diese Trennschärfe ließ eine unveränderte Verwendung der Kodierleitfäden in Studie S5 zu.

5.4.4 Abschluss der Studie S4

Die Zwischenstudie S4 diente schwerpunktmäßig der Implementierung kontextualisierter Fragestellungen und der Überprüfung der Stabilität des Instruments sowie dessen Kontextunabhängigkeit.

Hinsichtlich der Kontextualisierung der offenen Fragestellungen konnte besonders im Hinblick auf epistemologische und naturwissenschaftliche Bewertungsfaktoren ein deutlich verbessertes Antwortverhalten der Teilnehmenden beobachtet werden. Dies lässt den Schluss zu, dass ungerichtete Fragestellungen zu beeinflussenden Bewertungsfaktoren die tatsächlich berücksichtigten Perspektiven nur ungenügend abdecken und daher konkrete Anforderungssituationen in Form von Kontexten notwendig sind.

In Bezug auf die Kontextunabhängigkeit konnten zwischen den beiden Gruppen, welche separat voneinander befragt wurden, keine nennenswerten Unterschiede hinsichtlich des Antwortverhaltens festgestellt werden. Auch die etablierten Items des CRT und SRIS wiesen keine bedeutsamen Abweichungen im Gruppenvergleich auf und zeigten vergleichbare Ergebnisse mit den Gruppen vorangegangener Studien. Die Antworten der offenen Fragestellungen waren vergleichbar ausgeprägt wie die geschlossenen Items, was sich in ähnlichen Verständnisstrukturen der Begrifflichkeiten und hinzugezogenen Perspektiven innerhalb der Bewertungskontexte ausdrückte. Obwohl in den offenen Fragen der Fragebögen unterschiedliche Kontexte eingeführt wurden, ließ die statistische Auswertung mittels des Mann-Whitney-U-Tests keine signifikanten Unterschiede zwischen den beiden Gruppen hinsichtlich der verwendeten Bewertungsperspektiven erkennen. Trotz dieser übergeordneten Kontextunabhängigkeit konnten auf qualitativer Ebene leichte Unterschiede hinsichtlich der Gewichtung einzelner Bewertungsaspekte beobachtet werden, was sich jedoch nicht auf das untersuchte Konstrukt der Denkperspektiven im Bewertungsprozess auswirkte. Dieses Ergebnis deutet darauf hin, dass die zu untersuchenden Fragestellungen sowie die zugrunde liegenden Denkprozesse weitgehend kontextunabhängig sind und nicht signifikant durch die Variation des zu bewertenden Settings beeinflusst werden. Gleichzeitig könnte die Stabilität des Instruments als Indiz für die Reliabilität interpretiert werden, da insbesondere die konsistente Messung über verschiedene Kontexte und Gruppen hinweg darauf hindeutet, dass das Instrument

stabil die relevanten Konstrukte erfasst. Überdies könnte diese Kontextunabhängigkeit auch die Konstruktvalidität stärken, da sie nahelegt, dass die Fragestellung das Konstrukt unabhängig von äußeren Einflüssen misst. Trotzdem sind diese Ergebnisse aufgrund der geringen Stichproben- und Kontextvariabilität unter Vorbehalt zu interpretieren und bedürfen weiterer Überprüfungen durch nachfolgende Studien. Auf den Erkenntnissen von Studie S4 aufbauend wurden daher in der zweiten Hauptstudie S5 die bisher verwendeten Kontexte A (Solarenergie vs. Windkraft) und B (E-Mobilität vs. Brennstoffzellenantrieb) um die Kontexte C (Plastik), D (Fracking) und E (Stickoxidgrenzwerte) ergänzt, um eine vertiefte Überprüfung der Kontextunabhängigkeit vornehmen zu können.

5.5 Studie S5

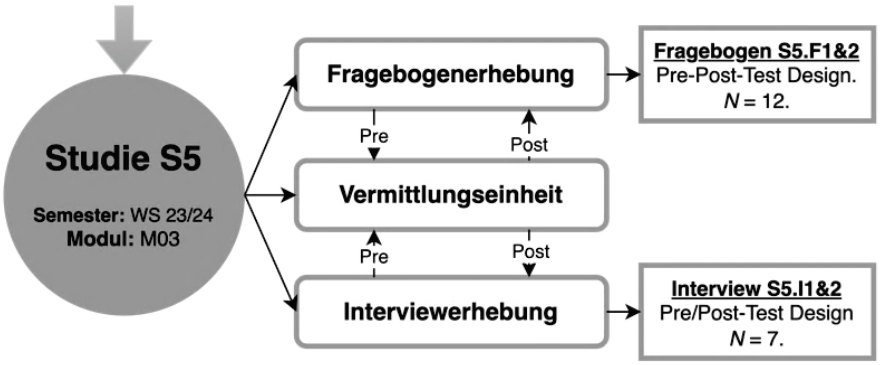


Abbildung 26: Übersicht über den Aufbau und die Erhebungsparameter der Studie S5

Die abschließende Hauptstudie S5 wurde im Wintersemester 2023/2024 durchgeführt. Zur weiteren Überprüfung der Stabilität des Instruments und zur Untersuchung einer fachlich fortgeschritteneren Stichprobe wurden die Erhebungen im Masterstudiengang Lehramt Chemie im Seminar *Vertiefende Aspekte der Chemiedidaktik (M1 HR/SP/GG)* durchgeführt. Im Rahmen dieses von einem anderen Dozierenden der Chemiedidaktik konzipierten Seminars übernahm der Untersuchende die Rolle eines Gastdozenten. In der Pre-Post-Fragebogenerhebung konnten insgesamt $N = 12$ vollständige Datensätze gewonnen werden, während an der Pre-Post-Interviewerhebung $N = 7$ Personen beteiligt waren (Abbildung 26). Da es sich um das Seminar eines anderen Dozierenden handelte, war der für die inhaltliche Vermittlung und die Durchführung der Systemisch-reflexiven Stellungnahme vorgesehene Zeitrahmen auf fünf Seminarsitzungen à 1,5 Stunden begrenzt. Ein Überblick über die Inhalte der Seminarsitzungen ist dem Kapitel 4.2.5 zu entnehmen.

Auf Grundlage der vorangegangenen Studien wurde die Kombination aus Interview- und Fragebogenerhebung zur Fremd- und Selbsteinschätzung beibehalten.

5.5.1 Interviewerhebung S5.I1&2

Die Interviewerhebung in Studie S5.I1&2 entsprach in Aufbau und Struktur der Studie S3.I1&2. Auch die eingesetzten Interviewleitfäden und Kodierleitfäden blieben unverändert. Zur weiterführenden Überprüfung der Kontextabhängigkeit wurden jedoch die Kontexte angepasst, wobei die Kontexte C (Plastik) und D (Fracking) als Bewertungssettings eingeführt wurden. Die Kontexte wurden für das Pre-Interview randomisiert und gleichmäßig auf die Teilnehmenden verteilt, wobei im Post-Interview der jeweils andere Kontext behandelt wurde. Der exemplarische Interviewleitfaden ist im Anhang (A.1.1) zu finden. Die verwendeten Kodierleitfäden sind ebenfalls dem Anhang (A.3.1.1) beigefügt. Nähere Erläuterungen zum Aufbau und zur Auswertung der Interviews können in der ausführlichen Darstellung der Studie S3.I1&2 (vgl. Kap. 5.3.1) nachgelesen werden.

Bei der Bewertung des Kontextes nannten die Teilnehmenden im Pre-Test im Durchschnitt 2,9 verschiedene *Aspekte der naturwissenschaftlichen Bewertung* (NW-A) als Grundlage für ihre Positionierung (Tabelle 111). Der mit etwa 86 % (6 von 7) größte Anteil der Interviewten sah die *Nachhaltigkeit* (NW-A-NA) als relevantes Kriterium für ihre Entscheidung an. Dies äußerte sich in gestellten Fragen, wie „Wie weit greift man da in die gegebenen Naturbestände ein? (S5.I1-WS23-D-M01-AB30VT, Pos. 36)“ oder „Wie betroffen ist die Natur? Wie betroffen ist […] Pflanzenwelt und Tierwelt?“ (S5.I1-WS23-D-M01-IF12CG, Pos. 39). 57,1 % (4 von 7) berücksichtigten zudem die *Umsetzbarkeit* (NW-A-UM) der Technologien. Dies zeigte sich in Aussagen, wie „[W]enn ich mich irgendwozu positioniere, […] dann mache ich mir häufig sehr viel Gedanken darum, inwieweit wird das wirklich Realität werden“ (S5.I1-WS23-C-M01-VM14OL, Pos. 34), oder in der Hervorhebung der Umsetzbarkeit als „ein sehr großer Aspekt“ (S5.I1-WS23-C-M01-MB14BO, Pos. 32). Aspekte, die sich auf die Kategorien *Effizienz* (NW-A-EF), *Kostenabwägung* (NW-A-KA) und *Rohstoffherkunft* (NW-A-RH) bezogen, wurden jeweils von nur einer Person genannt.

Im Post-Test erweiterten sechs der sieben Teilnehmenden (85,7 %) ihre Bewertungsgrundlage, was zum Anstieg der durchschnittlichen Anzahl berücksichtigter Perspektiven auf 4,3 führte. Neben den bereits im Pre-Test häufig benannten Kategorien *Nachhaltigkeit* (NW-A-NA) und *Umsetzbarkeit* (NW-A-UM) stieg die Berücksichtigung der Aspekte *Notwendigkeit* (NW-A-NO) (von 42,9 % auf 85,7 %) und *Kostenabwägung* (NW-A-KA) (von 14,3 % auf 71,4 %) deutlich an. Auch die *Effizienz* (NW-A) wurde häufiger (42,9 %, 3 von 7) als relevanter Aspekt für die Entscheidung angeführt.

Tabelle 111: S5.I1&2 - Ergebnisse der Interviewerhebung, *naturwissenschaftliche Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (NW), *N* = 7

Kat.	1. UK	2. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
NW: Naturwissenschaftliches Wissen	NW-A: Aspekte der naturwissenschaftlichen Bewertung	NW-A-KB: Keine Bewertung	0	0	0	0
		NW-A-EF: Effizenz	1	3	2	3
		NW-A-KA: Kostenabwägung	1	5	1	12
		NW-A-NA: Auswirkungen & Nachhaltigkeit	6	7	18	39
		NW-A-FT: Funktionsweise	4	2	8	4
		NW-A-UM: Umsetzbarkeit	4	5	13	11
		NW-A-RH: Rohstoffherkunft	1	2	1	2
		NW-A-NO: Notwendigkeit	3	6	6	22
Summe NW-A					**49**	**93**
Einzelzählung NW-A			**20**	**30**		

Die Kodierung der im Pre-Test berücksichtigten *selbstreflexiven Aspekte* (SR-A) ergab im Mittel 2,9 verschiedene *interne Faktoren* (SR-A-IF) und 1,7 unterschiedliche *externe Faktoren* (SR-A-EF) (Tabelle 112). Die am häufigsten (57 %, 4 von 7) benannten *internen Faktoren* (SR-A-IF) waren die Reflexion des eigenen *Fachwissens* (SR-A-IF-FW) und die Berücksichtigung der *persönlichen Relevanz* (SR-A-PR). Letztere spiegelte sich in Aussagen, wie „Natürlich, ob das Thema für mich relevant ist“ (S5.I1-WS23-D-M01-SJ21BO, Pos. 51), explizit wider. Die Aspekte *Emotionen* (SR-A-IF-EM), *Erfahrungen* (SR-A-IF-EF) und *Interesse* (SR-A-IF-IN) wurden von je drei Personen reflektiert.

Unter den *externen Faktoren* (SR-A-EF) wurde vor allem die *soziale Norm* (SR-A-EF-SN) als großer Einflussfaktor von 57,1 % (4 von 7) der Befragten benannt. Jeweils drei Personen empfanden zudem ihre eigene *Bildung* (SR-A-EF-BI) und die *Medien* (SR-A-EF-ME) als beeinflussend im Bewertungsprozess. Eine teilnehmende Person benannte die Auswirkungen von Medien mit der Aussage „Medien prägen mich sehr“ (S5.I1-WS23-D-M01-VF42IB, Pos. 56) kurz, aber direkt. Die Faktoren *Politik* (SR-A-EF-PO) und *Religion* (SR-A-EF-RE) spielten im Pre-Test nur für jeweils eine Person eine Rolle.

Tabelle 112: S5.I1&2 - Ergebnisse der Interviewerhebung, *selbstreflexive Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (SR), $N = 7$

Kat.	1. UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
SR: Selbstreflexion	SR-A: Aspekte der Selbstreflexion	SR-A-KS: Keine Selbstreflexion		0	0	0	0
		SR-A-IF: Interne Faktoren & introspektive Überprüfung	SR-A-IF-FW: Fachwissen	4	6	5	10
			SR-A-IF-PR: Persönliche Relevanz	4	6	7	24
			SR-A-IF-MW: Moral und Werte	0	4	0	7
			SR-A-IF-HA: Handlungen	1	3	1	4
			SR-A-IF-EM: Emotionen	3	5	4	15
			SR-A-IF-EF: Erfahrungen	3	5	5	15
			SR-A-IF-IN: Interesse	3	4	6	5
			SR-A-IF-BV: Befangenheit, Vorurteile	2	5	4	11
			SR-A-IF-PK: Persönlichkeit	0	4	0	4
			SR-A-IF-KV: Kognitive Verzerrungen	0	2	0	3
		Summe SR-A-IF				**32**	**98**
		Einzelzählung SR-A-IF		**20**	**44**		
		SR-A-EF: Bewusstsein für externe und interne Einflussfaktoren	SR-A-EF-SN: Soziale Norm/Umfeld	4	7	12	44
			SR-A-EF-ME: Medien	3	6	5	17
			SR-A-EF-BI: Bildung	3	6	6	13
			SR-A-EF-PO: Politik	1	3	3	8
			SR-A-EF-RE: Religion	1	1	1	1
		Summe SR-A-EF				**27**	**83**
		Einzelzählung SR-A-EF		**12**	**23**		

Im Post-Test stieg die Anzahl differenter Bewertungsperspektiven mit *selbstreflexivem Bezug* (SR-A) bei jeder Person stark an. Durchschnittlich benannten die Befragten im zweiten Interview 6,3 verschiedene *interne Faktoren* (SR-A-IF). Fast alle Teilnehmen-

den (85,7 %, 6 von 7) griffen hier die Kategorien *Fachwissen* (SR-A-IF-EF) und *Persönliche Relevanz* (SR-A-IF-PR) auf und jeweils 71,4 % (5 von 7) reflektierten die Einflüsse ihrer *Emotionen* (SR-A-IF-EM), *Erfahrungen* (SR-A-IF-EF) und ihrer persönlichen *Befangenheit* (SR-A-IF-BV). Diese Reflexion der Befangenheit drückte sich aus in Aussagen, wie „Mir ist halt bewusst, dass wenn ich etwas lese, was meine eigene Meinung besitzt, dass ich dem mehr Wertung zugebe" (S5.I2-WS23-D-M01-MB14BO, Pos. 36), oder in Fragestellungen, wie „Habe ich mich jetzt nur einseitig informiert? Habe ich wirklich nur die Sachen informiert, wo ich genau weiß […] das geht, glaube ich, in die Richtung, die ich mir eh schon vorgedacht habe?" (S5.I2-WS23-C-M01-AB30VT, Pos. 36). Auch die Kategorien *Persönlichkeit* (SR-A-IF-PK) sowie *Moral und Werte* (SR-A-IF-MW), die vorher nicht berücksichtigt wurden, wurden im Post-Test von 57,1 % (4 von 7) der Teilnehmenden aufgegriffen und zum Teil sehr klar als Einflussfaktoren benannt, wie folgendes Beispiel zeigt: „Das sind eigentlich meine höheren Ziele und meine Werte und meine Moral, die hinter meiner Meinung steckt [sic]" (S5.I2-WS23-D-M01-NB13SH, Pos. 40). Zwei Personen reflektierten zudem gezielt *kognitive Verzerrungen* (SR-A-IF-KV) in Aussagen, wie „[W]ie stark dann Verzerrungen entstehen, um dann zu überlegen, okay, welche Perspektive habe ich denn jetzt eigentlich eingenommen?" (S5.I2-WS23-C-M01-AB30VT, Pos. 56).

Auch die Menge der unterschiedlichen *externen Faktoren* (SR-A-EF), die im Post-Interview aufgegriffen wurden, verdoppelte sich fast von 1,7 auf 3,3. Alle Teilnehmenden benannten dabei ihr *soziales Umfeld* (SR-A-EF-SN) als prägend für ihre Positionierung und die Nennungen der Kategorien *Medien* (SR-A-EF-ME) und *Bildung* (SR-A-EF-BI) stiegen auf jeweils 85,7 % (6 von 7). Als Beispiel für Aussagen, die sich auf den Einfluss der Bildung bezogen, dient folgendes Zitat: „Also zum einen bin ich ja in einer recht privilegierten Situation, dass ich studieren kann. Das können viele Menschen nicht. Das heißt, ich habe Zugang zu ganz anderen Arten und Weisen der Informationsbeschaffung, die jetzt andere vielleicht nicht wissen" (S5.I2-WS23-D-M01-MB14BO, Pos. 42). Auch der potenzielle Einfluss der *Politik* (SR-A-EF-PO) wurde von drei Personen (42,9 %) aufgegriffen.

Die Untersuchung der *epistemologischen Aspekte* (EP-A) im Pre-Interview zeigte durchschnittlich 3,3 verschiedene Kriterien zur Bewertung von Quellen (Tabelle 113). Dabei standen primär die *Herkunft* (EP-A-B-HK) der Informationen sowie die *Übereinstimmung mit anderen Quellen* (EP-A-B-KO) bei jeweils 71,4 % der Befragten (5 von 7) im Vordergrund. Diese Art der *Konsistenzprüfung* (EP-A-B-KO) fand sich in Zitaten, wie „Dann würde ich […] versuchen andere Quellen zu finden, die den Artikel unterstützen oder vielleicht auch widerlegen" (S5.I1-WS23-D-M01-AB30VT, Pos. 48),

wieder. Auch die *Transparenz* (EP-A-B-TR), potenzielle *Interessenkonflikte* (EP-A-B-IK) oder die *Aktualität von Informationen* (EP-A-B-AK) spielten für jeweils 42,9 % (3 von 7) eine Rolle, wobei sich letztere in Fragen, wie „[W]as ist jetzt unser aktueller Forschungsstand [?]“ (S5.I1-WS23-C-M01-MB14BO, Pos. 30), ausdrückte. Die *Expertise* (EP-A-B-EX), *Präsentation* (EP-A-B-PR) oder *Methodik* (EP-A-B-ME) der Informationsquelle spielten für den Großteil der Teilnehmenden in der Erstbefragung keine Rolle.

Im Post-Test zeigten sich deutlich vermehrte Bewertungsperspektiven pro Proband*in, was sich in einer Mittelwerterhöhung von 3,3 auf 5,6 äußerte. Alle Teilnehmenden empfanden die *Herkunft* (EP-A-B-HK) und mögliche *Interessenkonflikte* (EP-A-B-IK) der Quellen als wichtige Kriterien, was in konkreten Fragen, wie „Wie werden die gefördert? Sind die unabhängig? Steht dahinter noch irgendeine Lobby oder sonst irgendwas?“ (S5.I2-WS23-C-M01-AB30VT, Pos. 26), deutlich wurde. Ebenfalls gewann die *Expertise* (EP-A-B-EX) an Berücksichtigung (71,5 % 5 von 7) und auch die *Methodik* (EP-A-B-ME) nahm an Bedeutung zu (57,1 %, 4 von 7). Die Berücksichtigung des methodischen Vorgehens lässt sich anhand des folgenden Beispielzitats ausführen: „Und wenn ich so eine Studie finde, dann würde ich mir als allererstes das Abstract durchlesen und dann auch schauen, wie jetzt die Studie abgelaufen ist. Also was haben die untersucht? Wie haben sie es untersucht?“ (S5.I2-WS23-D-M01-MB14BO, Pos. 16). Zwei Personen reflektierten darüber hinaus die *soziokulturelle Eingebundenheit* (EP-A-B-SE) wissenschaftlicher Erkenntnisse. Exemplarisch können hier folgende Zitate angeführt werden:

> „Und dabei ist mir aber erst mal aufgefallen, dass es ja gar nicht nur um meine Religion geht, sondern vielleicht auch um die derjenigen, die die Paper geschrieben haben, die ich lese“ (S5.I2-WS23-D-M01-IF12CG, Pos. 32).

> „Also, man müsste dann auch gucken, wo wurde die Studie durchgeführt? Wenn er [sic] zum Beispiel jetzt auch ein Land ist, was jetzt sehr groß in Fracking ist, denke ich, wird das auch noch mal Einfluss darauf haben, was vielleicht auch in den Studien dann kommuniziert werden darf. Und halt auch wie, dass die Regierung generell so ist mit Zensur und sowas. Wenn ich nur an China denke, da werden auch häufig Studien oder sowas […] zensiert“ (S5.I2-WS23-D-M01-MB14BO, Pos. 38).

Beide Zitate verdeutlichen ein Bewusstsein für die Einflussnahme gesellschaftlicher Strukturen auf die Veröffentlichung wissenschaftlicher Erkenntnisse.

Tabelle 113: S5.I1&2 - Ergebnisse der Interviewerhebung, *epistemologische Aspekte* bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug (EP), $N = 7$

Kat.	1. UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
EP: Epistemologie	EP-A: Aspekte der Epistemologie	EP-A-KB: Keine Bewertung der Quelle		0	0	0	0
		EP-A-B: Bewertung auf Basis epistemologischer Aspekte	EP-A-B-EX: Expertise	2	5	6	17
			EP-A-B-AK: Aktualität	3	2	3	4
			EP-A-B-IK: Interessenskonflikte	3	7	5	18
			EP-A-B-PR: Präsentation	1	3	1	7
			EP-A-B-HK: Herkunft	5	7	12	18
			EP-A-B-ME: Methodik	1	4	2	7
			EP-A-B-KO: Konsistenzprüfung	5	5	7	9
			EP-A-B-TR: Transparenz	3	4	3	6
			EP-A-B-SE: Soziokulturelle Eingebundenheit	0	2	0	4
Summe EP-A-B						**39**	**90**
Einzelzählung EP-A-B				**23**	**39**		

Die Auswertung der offenen Frage zum *Verständnis des Begriffs der Selbstreflexion* (SR-V) zeigte im Pre-Test bereits informierte Sichtweisen, welche ein *explizites Teilverständnis* (SR-V-ET) erkennen ließen. Fast alle Befragten (85,7 %, 6 von 7) benannten mindestens einen der Teilbereiche der Grunddefinition. Deutlich wurde, dass die Selbstreflexion schwerpunktmäßig als das *Überdenken von Handlungen* (SR-V-ET-HA) verstanden wurde (71,4 %, 5 von 7) (Tabelle 114). Exemplarisch kann ein Zitat angeführt werden, welches Selbstreflexion auf die Art versteht, „dass man selber auch hinterfragt, war das jetzt richtig, was ich da getan habe, oder ist das jetzt gerechtfertigt gewesen?“ (S5.I1-WS23-D-M01-VF42IB, Pos. 68). Drei Personen (42,9 %) verstanden unter Selbstreflexion die Einsicht in die Gefühlswelt im Sinne einer *Introspektion* (SR-V-ET-IS) und nur eine Person inkludierte in ihrem Verständnis das *Bewusstsein für externe Prägungen* (SR-V-ET-EX). 42,9 % (3 von 7) zeigten zudem teilweise naive Sichtweisen des *impliziten Teilverständnisses* (SR-V-IT). Hier wurde von jeweils einer Person die Selbstreflexion als *abwägender Prozess* (SR-V-IT-AB), *Prozess der Selbstverbesserung* (SR-V-IT-SV), *Empathie* (SR-V-IT-EM) oder unkonkretes *Hinterfragen* (SR-V-IT-HF) verstanden.

Tabelle 114: S5.I1&2 - Ergebnisse der Interviewerhebung, *Verständnis Selbstreflexion* (SR-V), *N* = 7

Kat.	1. UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
SR: Selbstreflexion	SR-V: Verständnis Selbstreflexion	SR-V-ND: nicht definierbar		0	0	0	0	N=0
		SR-V-IT: Implizites Teilverständnis	SR-V-IT-AB: Abwägen	1	1	1	1	N=1
			SR-V-IT-SV: Selbstverbesserung /Kritik	1	0	2	0	
			SR-V-IT-EM: Empathie	1	0	1	0	
			SR-V-IT-HF: Hinterfragen	1	1	1	1	
		Summe SR-V-IT				**5**	**2**	
		Einzelzählung SR-V-IT		**4**	**2**			
		SR-V-ET: Explizites Teilverständnis	SR-V-ET-IS: Introspektion	3	7	5	16	N=2
			SR-V-ET-HA: Überdenken von Handlungen	5	6	9	13	
			SR-V-ET-EX: Bewusstsein für externe Prägungen	1	5	1	10	
		Summe SR-V-ET				**15**	**39**	
		Einzelzählung SR-V-ET		**9**	**18**			

Trotz des bereits guten Verständnisses des Begriffs der Selbstreflexion verbesserte und konkretisierte sich das Verständnis im Post-Test noch weiter. Hier erreichten alle Teilnehmenden Niveau 2 und benannten pro Person mindestens zwei Teilbereiche des *expliziten Teilverständnisses* (SR-V-ET). Alle Teilnehmenden zählten im zweiten Interview die *Introspektion* (SR-V-ET-IS) zur Selbstreflexion, was sich beispielsweise an folgender Aussage festmachen lässt: „Wenn ich aber zum Beispiel mich selbst reflektiere, dann würde ich ja gucken, was habe ich jetzt vielleicht gedacht? Oder was habe ich in dem Moment gefühlt?“ (S5.I2-WS23-D-M01-MB14BO, Pos. 50). 85,7 % (6 von 7) ergänzten diese Sichtweise um das *Überdenken von Handlungen* (SR-V-ET-HA). Das Bewusstsein für *externe Prägungen* (SR-V-ET-EX) wurde zuvor von nur einer Person angeführt, während es nun von fünf Personen als zugehörig zur Begriffsdefinition erkannt wurde. Naivere Ansichten verschwanden fast gänzlich, bis auf eine Person, die weiterhin Selbstreflexion als *abwägenden Prozess* (SR-V-IT-AB) wahrnahm, und eine Person, die es, ergänzend zu den informierten Sichtweisen, als *Hinterfragen* (SR-V-IT-HF) beschrieb.

Der Begriff des *Kritischen Denkens* (KD-V) zeigte sich hingegen im Pre-Test deutlich weniger sinnerfüllt und informiert. Nur drei der sieben Personen (42,9 %) erreichten Niveau 1 und damit *Teilbereiche der Grunddefinition* (KD-V-TG). Der größte Anteil der

Befragten (71,4 %) verblieb auf Niveau 0 und fasste unter Kritischem Denken „hauptsächlich dieses aktive Hinterfragen“ (S5.I1-WS23-C-M01-MB14BO, Pos. 54) (*Hinterfragen*, KD-V-NA-HF), ohne dies näher zu konkretisieren. Jeweils zwei Personen (28,6 %) verstanden unter Kritischem Denken den Prozess der *Selbstverbesserung* (KD-V-NA-SV) oder das *Tadeln anderer* (KD-V-NA-TK) (Tabelle 115).

Tabelle 115: S5.I1&2 - Ergebnisse der Interviewerhebung, *Verständnis Kritisches Denken* (KD-V), *N* = 7

Kat.	1. UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
KD: Kritisches Denken	KD-V: Verständnis von KD	KD-V-NA: naive Vorstellung von KD	KD-V-NA-ND: nicht definierbar	1	0	1	0	N=0
			KD-V-NA-HF: Hinterfragen	5	4	8	5	
			KD-V-NA-SK: Skepsis	0	1	0	1	
			KD-V-NA-SV: Selbstverbesserung	2	1	2	2	
			KD-V-NA-TK: Tadeln/Korrigieren	2	2	3	5	
			KD-V-NA-SD: Selbstständiges Denken	1	0	1	0	
		Summe KD-V-NA:				**15**	**13**	
		Einzelzählung KD-V-NA:		**11**	**8**			
		KD-V-TG: Teilbereich der Grund-definition	KD-V-TG-SR: Selbstreflexion	0	5	0	14	N=1
			KD-V-TG-WI: Wissen	3	6	4	10	
			KD-V-TG-EP: Epistemologisches Verständnis	0	2	0	3	
		Summe KD-V-TG:				**4**	**27**	
		Einzelzählung KD-V-TG:		**3**	**13**			
		KD-V-MT: Zusammenspiel mehrerer Teilbereiche		**0**	**5**	**0**	**5**	N=2

Im Post-Test waren hingegen deutlich informiertere Sichtweisen erkennbar. Alle Teilnehmenden erreichten mindestens Niveau 1 und fassten unter Kritischem Denken vor allem die Aneignung von *Wissen* (KD-V-TG-WI) (85,7 %, 6 von 7) und die *Selbstreflexion* (KD-V-TG-SR) (71,4 %, 5 von 7). Diese war beispielsweise in folgender Aussage erkennbar: „Also Kritisches Denken heißt nicht nur den Fakt hinterfragen, sondern vielleicht auch dann an den Punkt zu gelangen, sich selber zu hinterfragen. Das ist das, was ich dann auch mit der Selbstreflexion meine“ (S5.I2-WS23-C-M01-AB30VT, Pos. 56). Das *epistemologische Verständnis* (KD-V-TG-EP) wurde von zwei Personen als Teil des Kritischen Denkens berücksichtigt. Die informierten Ansichten wurden zum Teil von naiveren Vorstellungen, wie der Verwendung der

Phrase des *Hinterfragens* (KD-V-NA-HF), begleitet. Dennoch zeigte sich in den Ausführungen der Teilnehmenden ein umfassenderes Verständnis, was sich unter anderem darin ausdrückte, dass fünf Befragte Kritisches Denken als *Zusammenspiel seiner Teilbereiche* (KD-V-MT) begriffen und damit die höchste Niveaustufe (2) erreichten.

Die letzten drei Kategorien der *fachlichen Bewertung* (NW-FB), der *Selbsteinschätzung des Fachwissens* (NW-SF) und der Einschätzung der *Fähigkeit zur Bewertung* (FB) wurden analog zur Studie S3.I1&2 kombiniert ausgewertet. Die Studierenden schätzten ihr Wissen zu den gestellten Kontexten größtenteils als *grundlegend* (NW-SF-GL) ein (85,7 %, 6 von 7), nur zwei Personen (28,6 %) sahen ihr Wissen als *unzureichend* (NW-SF-UZ) an (Tabelle 116). Sechs der sieben Befragten (85,7 %) gaben ebenfalls Informationen zum Kontext ohne verortbare Quelle auf Basis von Hörensagen an. Im Post-Test war die Einschätzung bezüglich des *Fachwissens* (NW-SF) vergleichbar zum Pre-Test.

Die *Bewertungsfähigkeit* (FB) wurde andererseits in beiden Interviews eher schwach eingeschätzt. Vier der sieben interviewten Personen (57,1 %) des Pre-Interviews und drei des Post-Interviews gaben an, sich *nicht in der Lage* (FB-NL) zu fühlen, eine Bewertung vorzunehmen (Tabelle 117). Bei zwei Teilnehmenden war im Post-Interview keine Kodierung möglich, da keine Aussagen mit ausreichendem inhaltlichem Bezug zur analysierten Kategorie vorlagen.

Tabelle 116: S5.I1&2 - Ergebnisse der Interviewerhebung, *Selbsteinschätzung Fachwissen (NW-SF)* und *Fachwissenschaftliche Bewertung (NW-FB*), *N* = 7

Kat.	1. UK	2. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
NW: Naturwissenschaf	NW-SF: Selbsteinschätzung Fachwissen	NW-SF-UZ: Unzureichend	2	3	4	5	**N=0**
		NW-SF-GL: Grundlegend	6	3	7	5	**N=1**
		NW-SF-UR: Umfangreich	0	1	0	1	**N=2**
NW: Naturwissenschaftliche	NW-FB: Fachwissenschaftliche Bewertung	NW-FB-NV: Nicht vorhanden	0	0	0	0	**N=0**
		NW-FB-UG: Ungeprüft/ Unbelegt	6	7	12	12	**N=1**
		NW-FB-FU: Fundiert	1	0	1	0	**N=2**

Tabelle 117: S3.I1&2 - Ergebnisse der Interviewerhebung, *Fähigkeit zur Bewertung (FB)*, *N* = 7

Kat.	1. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
FB: Fähigkeit zur Bewertung	FB-NL: Nicht in der Lage	4	3	4	3	**N=0**
	FB-UE: Unter Einschränkungen in der Lage	3	1	3	1	**N=1**
	FB-IL: In der Lage	1	1	1	1	**N=3**
Summe FB-UE/IL:				**4**	**2**	
Einzelzählung FB-UE/IL:		**4**	**2**			

Die kombinierte Betrachtung zeigte, dass bei einer teilnehmenden Person (NB13SH) im Pre-Interview eine Diskrepanz zwischen der *Selbsteinschätzung des Fachwissens* (NW-SF) und der Einschätzung der *Fähigkeit zur Bewertung* (FB) bestand. So schätzte sie ihr Fachwissen auf Nachfrage mit den Worten „Zu wenig, würde ich behaupten" (S5.I1-WS23-C-M01-NB13SH, Pos. 24) ein, sah sich aber im späteren Interviewverlauf zu einer Positionierung *in der Lage* (FB-IL). Gleichzeitig gehörte diese Person zu den Teilnehmenden, die im Pre-Interview am wenigsten *naturwissenschaftliche Aspekte* (NW-A) erfragten, und berücksichtigte lediglich die *Umsetzbarkeit* (NW-A-UM) und die *Notwendigkeit* (NW-A-NO). Dies legt eine subjektive Bewertungsfähigkeit nahe, welche in Ermangelung fachlicher Perspektiven vermutlich auf affektive Faktoren zurückzuführen ist, wie beispielsweise persönliche Überzeugungen, allgemeine Werte oder intuitive Einschätzungen.

Eine direkte Gegenüberstellung der Ergebnisse ist in Tabelle 118 dargestellt. Diese zeigt analog zur Darstellung in Studie S3.I1&2 die Anzahl der unterschiedlichen Aspekte pro Proband*in und pro *Teilbereich* (NW-A, EP-A und SR-A). Die beiden Subkategorien der *internen* (SR-A-IF) und *externen Faktoren* (SR-A-EF) sind separat ausgewiesen. Für die Begriffsdefinitionen wurde die Anzahl der unterschiedlichen Teilbereiche erfasst, die eine informierte Definition abbilden. Beim *Verständnis der Selbstreflexion* (SR-V) schließt dies das *explizite Teilverständnis* (SR-V-ET) ein, während es beim Verständnis des *Kritischen Denkens* (KD-V) die *Teilbereiche der Grunddefinition* (KD-V-TG) beinhaltet.

Tabelle 118: S5.I1&2 - Ergebnisse der Interviewerhebung, Übersicht über die Einzelzählungen der verschiedenen Teilbereiche, $N = 7$

Einzel-zählung	NW-A		SR-A		EP-A		KD-V		SR-V		SR-A-IF		SR-A-EF	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
AB30VT	4	6	4	5	6	10	0	2	0	2	3	7	3	3
IF12CG	3	3	4	7	4	11	0	2	1	3	2	7	2	4
MB14BO	2	5	5	7	2	11	0	1	1	3	1	7	1	4
NB13SH	2	4	3	4	4	11	1	3	2	2	3	7	1	4
SJ21BO	3	4	2	4	4	7	1	2	1	2	3	4	1	3
VF42IB	3	4	2	5	6	9	0	1	1	3	5	7	1	2
VM14OL	3	4	3	7	6	8	1	2	3	3	3	5	3	3
MW	2,600	4,100	4,200	7,300	0,500	3,800	0,200	2,300	0,900	1,800	2,500	5,400	1,700	1,900
Stabw.	0,966	1,853	1,814	2,163	0,527	2,044	0,422	0,823	0,876	0,632	1,434	1,647	1,059	0,994
Varianz	0,933	3,433	3,289	4,678	0,278	4,178	0,178	0,678	0,767	0,400	2,056	2,711	1,122	0,989
Max	4	6	8	11	1	7	1	3	2	3	5	8	3	3
Min	1	1	1	5	0	1	0	1	0	1	1	3	0	0
Median	3	5	4	7	0,5	3,5	0	2,5	1	2	2	5	1,5	2
Modus	3	5	4	5	0	3	0	3	0	2	2	5	1	2

Auf Grundlage dieser Daten wurde der Wilcoxon-Test als Signifikanztest durchgeführt (Tabelle 119).

Tabelle 119: S5.I1&2 - Ergebnisse der Interviewerhebung, Wilcoxon-Test der Einzelzählungen der verschiedenen Teilbereiche, $N = 7$

S5.I1&2 - Wilcoxon-Test

Teststatistikena							
Einfachzählung	NW_A_Post - NW_A_Pre	SR_A_Post - SR_A_Pre	EP_A_Post - EP_A_Pre	KD_V_Post - KD_V_Pre	SR_V_Post - SR_V_Pre	SR_A_IF_Post - SR_A_IF_Pre	SR_A_EF_Post - SR_A_EF_Pre
Z	-2,232b	-2,375b	-2,379b	-2,428b	-2,121b	-2,375b	-2,041b
Asymp. Sig. (2-seitig)	0,026	0,017	0,018	0,015	0,034	0,018	0,041
Exakte Sig. (2-seitig)	0,031	0,016	0,016	0,016	0,063	0,016	0,063
Exakte Sig. (1-seitig)	0,016	0,008	0,008	0,008	0,031	0,008	0,031
Punkt-Wahrscheinlichkeit	0,016	0,008	0,008	0,008	0,031	0,008	0,031
a. Wilcoxon-Test							
b. Basiert auf negativen Rängen.							
Effektstärke	0,597	0,636	0,635	0,649	0,567	0,635	0,545

Auch hier ließen sich, vergleichbar mit den Ergebnissen der Interviewerhebung der Studie S3.I1&2, bei einem Signifikanzniveau $\alpha = 0,05$ statistisch signifikante Unterschiede in allen Teilbereichen erkennen. Sowohl im Teilbereich des *naturwissenschaftlichen Wissens* (NW-A) ($z = -2,232$, $p = 0,031$, $r = 0,596$) als auch in den Teilbereichen der *Selbstreflexion* (SR-A) ($z = -2,375$, $p = 0,016$, $r = 0,635$) und des *epistemologischen*

Verständnisses (EP-A) (z = -2,379, p = 0,016, r = 0,636) wurden signifikante Veränderungen mit hoher Effektstärke hinsichtlich der bei der Bewertung hinzugezogenen Perspektiven erkennbar. Die größte Veränderung im Teilbereich der Selbstreflexion war dabei bei den *internen Faktoren* (SR-A-IF) zu verorten.

Das *Verständnis des Begriffs der Selbstreflexion* (SR-V) lag bereits im Pre-Test auf einem informierten Niveau vor und zeigte im Pre-Post-Vergleich keine signifikanten Veränderungen, auch wenn die exakte Signifikanz von p = 0,063 und einer Teststatistik von z = -2,121 Veränderungstendenzen in eine positive Richtung erkennen lässt. Das *Verständnis vom Kritischen Denken* (KD-V) hingegen veränderte sich deutlich signifikant mit einer hohen Effektstärke (z = -2,428, p = 0,016, r = 0,649) in Richtung eines informierteren Verständnisses.

Diese Veränderungen wurden in Abbildung 27 grafisch für die gesamte Stichprobe mit Bezug auf das Triadische Modell Kritischen Denkens dargestellt. Der Schwerpunkt der Darstellung lag wie in Studie S3 (vgl. Kap. 5.3.1.2) auf den Bewertungsteilbereichen des *naturwissenschaftlichen Wissens* (NW-A), des *epistemologischen Verständnisses* (EP-A) und der *Selbstreflexion* (SR-A) sowie auf den *Begriffsverständnissen von Kritischem Denken* (KD-V) und *Selbstreflexion* (SR-V).

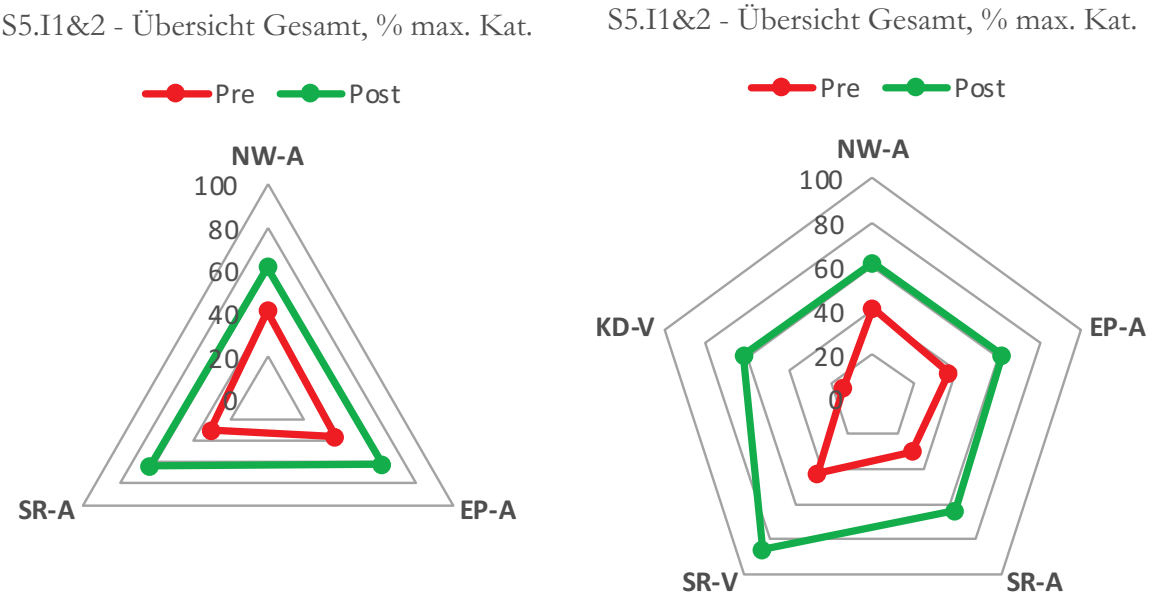


Abbildung 27: S5.I1&2 - Gesamtübersicht über die prozentualen Anteile an den einzelnen Teilbereichen (% max. Kat. bezeichnet den prozentualen Anteil an der maximalen Anzahl möglicher Kategorien)

In Abbildung 28 sind zudem exemplarische Vergleichsdiagramme einzelner Proband*innen zur Veranschaulichung der individuellen Entwicklung abgebildet.

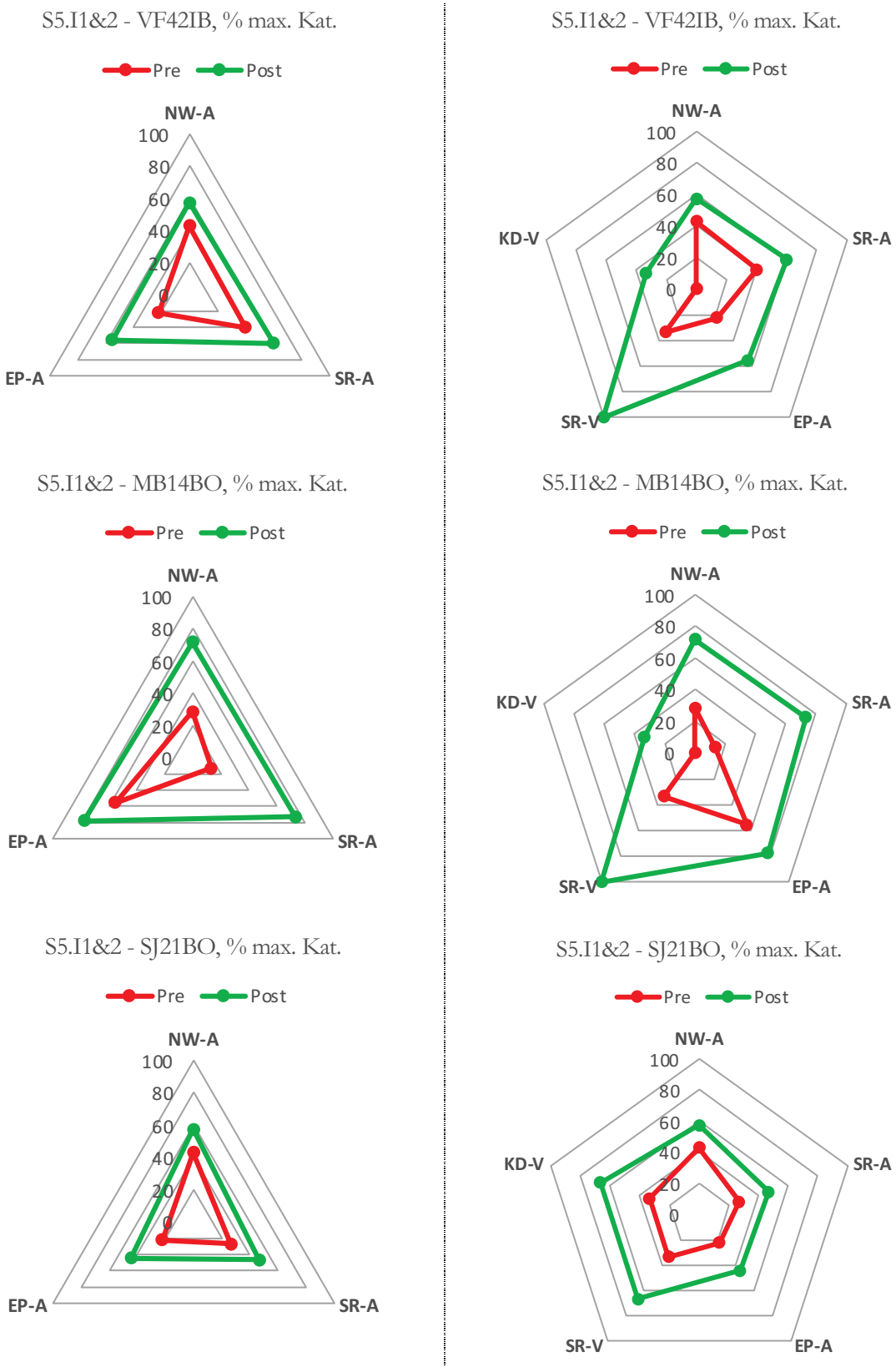


Abbildung 28: S5.I1&2 - Exemplarische Auswahl individueller Vergleichsdiagramme
(% max. Kat. bezeichnet den prozentualen Anteil an den maximal möglichen Kategorien)

Auch in dieser Studie wurden, zusätzlich zu den Pre-Post-Interviews, ergänzende Follow-Up-Interviews mit fünf der sieben Teilnehmenden durchgeführt, welche im

Abstand von sechs Monaten stattfanden. Da die Ergebnisse der Studie S4 eine Kontextunabhängigkeit des Instruments nahelegten und für eine erneute Befragung ein unbekannter Kontext notwendig war, wurde der neue Kontext E (Stickoxidgrenzwerte) als Setting eingesetzt (s. auch Kap. 4.2.6.6).

Übereinstimmend mit den Ergebnissen der Erhebung S3.I1&2 zeigten sich die Resultate der Follow-Up-Interviews vergleichbar zu denen der Post-Befragungen (Abbildung 29). Die Anzahl an zur Bewertung hinzugezogenen Aspekten blieb in den meisten Fällen konstant hoch und erweiterte sich in einigen Fällen sogar. Auch das Verständnis der einzelnen Begrifflichkeiten blieb erhalten und in allen Fällen auf dem höchsten Niveau (2).

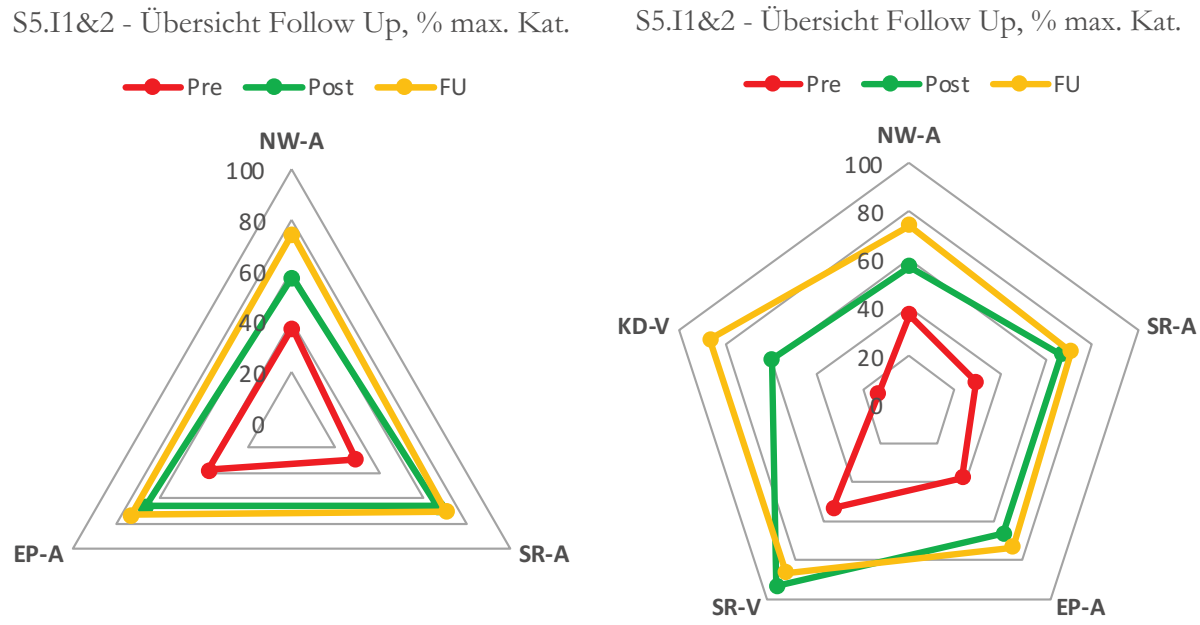


Abbildung 29: S5.I1&2 + Follow-Up - Gesamtübersicht über die prozentualen Anteile an den einzelnen Teilbereichen (% max. Kat. bezeichnet den prozentualen Anteil an der maximalen Anzahl möglicher Kategorien)

Abbildung 30 zeigt ergänzend drei exemplarische Vergleichsdiagramme von Teilnehmenden der Follow-Up-Erhebung:

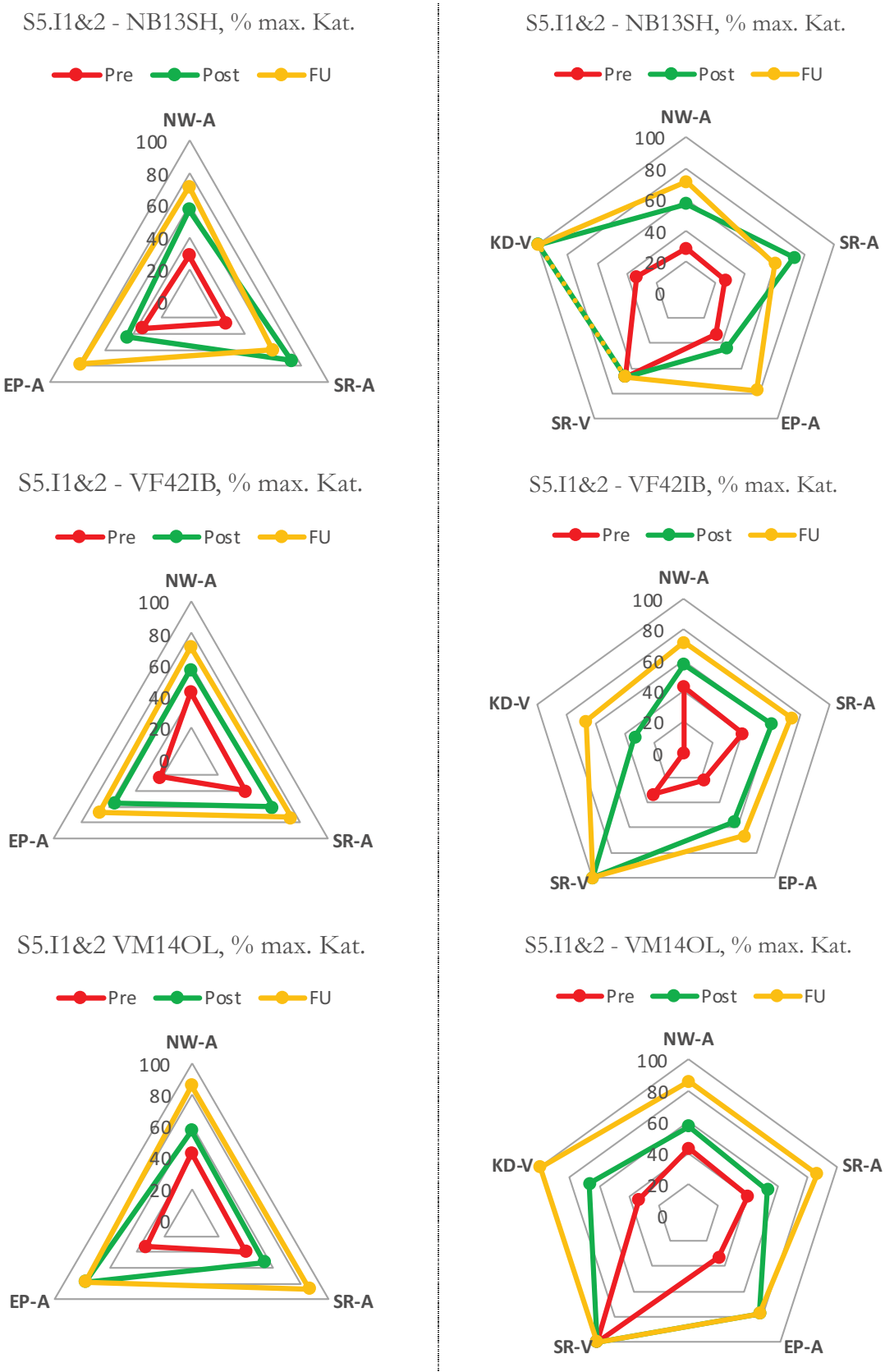


Abbildung 30: S5.I1&2 + FU - S5.I1&2 - Exemplarische Auswahl individueller Vergleichsdiagramme (% max. Kat. bezeichnet den prozentualen Anteil an der maximalen Anzahl möglicher Kategorien)

Auch im Anschluss an die Follow-Up-Interviews der Studie S5 wurden, angesichts des positiven Verlaufs der Interviews, gezielte Nachfragen zu den reflexiven Prozessen und den wahrgenommenen Langzeitwirkungen gestellt. Die Befragten führten insbesondere ihre im Alltag gesteigerte Sensibilität in Bewertungsprozessen als Grund für den zum Teil besseren Verlauf der Follow-Up-Interviews im Vergleich zum Post-Interview an. Eine Person brachte es im Interview folgendermaßen zum Ausdruck: „Ja, aber irgendwie schwingt bei mir immer so ein bisschen im Hinterkopf in letzter Zeit mit: Das hast du nicht überprüft“ (S5.3-WS23-E-M01-NB13SH, Pos. 40).

Auch die Resultate der Interviewerhebung S5.I1&2 weisen auf eine Nachhaltigkeit der erzielten Ergebnisse hin, welche sich bereits in Studie S3 angedeutet hatte. Diese stützen zum einen die Annahme der Funktionalität des Instruments zur Erhebung Kritischen Denkens im Sinne eines multiperspektivischen Denkansatzes, zum anderen festigen sie die Hypothese, dass Kritisches Denken, wie es hier definiert wird, grundsätzlich förderbar ist. Das Instrument erweist sich damit als valide in der Messung derjenigen Prozesse, die in der Vermittlungseinheit adressiert werden. Gleichzeitig zeigt sich, dass die Vermittlungseinheit gezielt jene Konstrukte beeinflusst, die das Instrument zu messen intendiert. Diese wechselseitige Beziehung zwischen Intervention und Messung unterstreicht die Passgenauigkeit des Instruments für die intendierten Bildungs- und Veränderungsprozesse und damit die Validität des Instruments.

5.5.2 Fragebogenerhebung

Der in der Erhebung S5.F1&2 eingesetzte Fragebogen entsprach in vollem Umfang den Aufgabenstellungen, der Struktur und den eingesetzten Kontexten des in Studie S4 verwendeten Fragebogens, der wiederum auf dem in Studie S3.F1&2 entwickelten Instrument basierte. Angesichts dessen wird zur Vermeidung von Redundanzen auf eine erneute Darstellung verzichtet. Der Aufbau und die Auswertung des Fragebogens sind in Kapitel 5.3.2 zur Studie S3.F1&2 aufgeführt, während die Ergänzungen aus Studie S4.F1&2 in Kapitel 5.4 erläutert werden.

Im Gegensatz zur Erhebung in Studie 4 wurde in der Erhebung S5.F1&2 ein Pre-Post-Test-Vergleich durchgeführt. Für die kontextualisierten Fragen wurde infolgedessen im Pre-Test der Kontext B (Elektromobilität vs. Brennstoffzelle) und im Post-Test der Kontext A (Solarenergie vs. Windkraft) eingesetzt. Die Zuordnung der Kontexte zu Pre- und Post-Test verlief zufallsbasiert.

Die Auswertung der Fragenblöcke erfolgte ebenfalls analog zu den Erhebungen S3.F1&2 sowie S4.F1&2. Zusätzliche Details zu den durchgeführten Rechenoperationen sind im Kapitel 4.3.2 zu finden.

Bei der Auswertung der offenen Frage (VF1) zum *Verständnis der Selbstreflexion* (VS) zeigte sich bereits im Pre-Test ein informiertes Bild bei den Teilnehmenden (Tabelle 120).

Tabelle 120: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis Selbstreflexion* (VS), *N* = 12

Kat.	1.UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
VS: Verständnis Selbstreflexion	**VS-ND: Nicht definierbar (ND)**			0	0	0	0	**N=0**
	VS-IT: Implizites Teilverständnis		VS-IT-AB: Abwägen	1	0	1	0	**N=1**
			VS-IT-SV: Selbstverbesserung	0	0	0	0	
			VS-IT-EP: Empathie	0	1	0	1	
			VS-IT-HF: Hinterfragen	1	0	1	0	
	Summe VS-IT:					**2**	**1**	
	Einzelzählung VS-IT:			**2**	**1**			
	VS-ET: Explizites Teilverständnis	**VS-ET-IS: Introspektion**	VS-ET-IS-AL: Introspektion allg.	0	0	0	0	**N=2**
			VS-ET-IS-PK: Persönlichkeit	0	1	0	1	
			VS-ET-IS-GE: Gedanken	5	5	5	5	
			VS-ET-IS-EM: Emotionen	0	4	0	4	
			VS-ET-IS-WI: Wissen	1	0	1	0	
			VS-ET-IS-AN: Ansichten/Meinung	3	6	3	6	
			VS-ET-IS-MW: Moral/Werte	0	3	0	3	
			VS-ET-IS-FA: Fähigkeiten	0	1	0	1	
			VS-ET-IS-EF: Erfahrungen	0	1	0	1	
			VS-ET-HA: Handlungen überdenken	9	8	9	9	
			VS-ET-EX: Externe Prägungen	1	2	1	2	
	Summe					**19**	**32**	
	Einzelzählung			**5**	**9**			

Diese verstanden Selbstreflexion schwerpunktmäßig (75 %, 9 von 12) als das *Überdenken von Handlungen* (VS-ET-HA). 42,7 % (5 von 12) fassten darunter die *Reflexion der eigenen Gedanken* (VS-ET-IS-GE), wohingegen 25 % (3 von 12) die *Reflexion der eigenen Ansichten* (VS-ET-IS-AN) zur Definition zählten.

Im Post-Test differenzierte sich das bereits informierte Verständnis aus dem Pre-Test weiter aus. Neben der *Reflexion der eigenen Gedanken* (VS-ET-IS-GE) und der *Reflexion der eigenen Ansichten* (VS-ET-IS-AN) traten hier verstärkt Bezüge zu *Emotionen* (VS-ET-IS-EM) (33,3 %, 4 von 12) und *Wertvorstellungen* (VS-ET-IS-MW) (25 %, 3 von 12) auf. Auch die Kategorien *Persönlichkeit* (VS-ET-IS-PK), *Fähigkeiten* (VS-ET-IS-FA) und *Erfahrungen* (VS-ET-IS-EF), welche im Pre-Test seitens der Befragten außer Acht gelassen wurden, konnten bei je einer Person kodiert werden.

Im Gegensatz zum Verständnis der Selbstreflexion lag das *Verständnis Kritischen Denkens* (VK) im Pre-Test deutlich weniger informiert vor. Der größte Anteil der Befragten (58,3 %, 7 von 12) verstand unter Kritischem Denken das *Hinterfragen von Dingen* (VK-NA-HF), ohne dies weiter auszuführen. Drei der zwölf Teilnehmenden (25 %) sahen zudem darin eine grundsätzlich *skeptische Haltung* (KV-NA-SK). Die Hälfte der Befragten verblieb damit auf dem niedrigsten Verständnisniveau (0). Die andere Hälfte erreichte Niveau 1 und damit ein informiertes *Teilverständnis* (VK-TG). Je zwei der Teilnehmenden (16,7 %) fassten unter der Begrifflichkeit eine *multiperspektivische Betrachtungsweise* (KV-TG-MP), die Integration von *Selbstreflexion* (KV-TG-SR) oder den *Umgang mit Wissen* (KV-TG-WI) (Tabelle 121).

Die Ergebnisse des Post-Tests ließen ein deutlich informierteres Verständnis Kritischen Denkens erkennen. Fast alle Befragten (91,7 %, 11 von 12) inkludierten die *Selbstreflexion* (KV-TG-SR) in ihr Verständnis. Die Hälfte subsumierte zudem die Anwendung *epistemologischen Verständnisses* (KV-TG-EP) zur Bewertung von Quellen unter dem Begriff des Kritischen Denkens und der Anteil der Personen, die die Reflexion von *Wissen* (KV-TG-WI) mit einbezogen, stieg auf 33,3 % (4 von 12). Die Phrase des *Hinterfragens* (KV-NA-HF) wurde hingegen lediglich einmal erwähnt. Alle Studierenden erreichten damit ein fortgeschrittenes Verständnisniveau (1), 83,3 % (10 von 12) sogar das höchste Verständnisniveau (2).

Tabelle 121: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis Kritisches Denken* (VK), *N* = 12

Kat.	1.UK	2. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
VK: Verständis Kritisches Denken	VK-NA: Naive Vorstellung	VK-NA-ND: Nicht definierbar	0	0	0	0	N=0
		VK-NA-HF: Hinterfragen	7	1	7	1	
		VK-NA-SK: Skepsis	3	0	3	0	
		VK-NA-SV: Selbstkritik/-verbesserung	0	0	0	0	
		VK-NA-TK: Tadeln/Kritisieren	0	0	0	0	
		VK-NA-SD: Selbstständiges Denken	1	0	1	0	
	Summe VK-NA:				**11**	**1**	
	Einzelzählung VK-NA:		**3**	**1**			
	VK-TG: Teilbereich der Grunddefinition	VK-TG-MP: Multiperspektivität	2	3	2	3	N=1
		VK-TG-SR: Selbstreflexion	2	11	2	13	
		VK-TG-WI: Wissen	2	4	2	4	
		VK-TG-EP: Epistemologisches Verständnis	1	6	1	7	
	Summe VK-TG:				**7**	**27**	
	Einzelzählung VK-TG:		**4**	**4**			
	VK-MT: Zusammenspiel mehrerer Teilbereiche		**1**	**10**			N=2

Im Pre-Fragebogen gaben nur zwei Personen (16,7 %) an, *keine Vorstellung* (KV-NA-UK) vom Terminus der *kognitiven Verzerrungen* (KV) zu haben (Tabelle 122). Der Großteil der Befragten fasste darunter *Denk- und Wahrnehmungsfehler* (KV-TG-DF) (66,7 %, 8 von 12) und *interne Einflüsse* (KV-TG-IF) (33,3 %, 4 von 12), was ein *informiertes Verständnis* (KV-TG) widerspiegelte.

Tabelle 122: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Verständnis kognitive Verzerrungen* (KV), *N* = 12

Kat.	1.UK	2. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post	Niveau
KV: Verständnis kognitive Verzerrungen	KV-NA: Naive Vorstellungen	KV-NA-UK: Unkonkret/Keine Vorstellung	2	0	11	0	N=0
		KV-NA-FV: Fehlvorstellungen	0	0	1	0	
		Summe NA			**12**	**0**	
		Einzelzählung NA	**1**	**0**			
	KV-TG: Teilbereich der Grunddefinition	KV-TG-DF: Denk-und Wahrnehmungsfehler	8	10	3	11	N=1
		KV-TG-IF: Einfluss von internen Faktoren	4	4	1	4	
		KV-TG-EF: Einfluss von externen Faktoren	0	5	0	5	
	Summe TG				**4**	**20**	
	Einzelzählung TG		**2**	**3**			

Im Post-Test wiesen alle Befragten konkrete Vorstellungen des Begriffs der kognitiven Verzerrungen auf und 20 % (5 von 12) fassten nun auch die *Einflüsse externer Faktoren* (KV-TG-EF) darunter. Zudem differenzierten sich die Vorstellungen dahingehend aus, dass die Hälfte der Studierenden mehr als einen *Teilbereich der Grunddefinition* (KV-TG) in ihr Verständnis einbezog.

Die Auswertung der offenen Fragen zu den *Bewertungsfaktoren* (BF) zeigte im Pre-Test auf *selbstreflexiver Ebene* (BF-SR) bei allen Teilnehmenden eine Reflexion des eigenen *Fachwissens* (BF-SR-IF-FW). 83,3 % (10 von 12) benannten zusätzlich die *persönliche Relevanz* (BF-SR-IF-PR) als wichtigen Faktor im Bewertungsprozess. 41,7 % (5 von 12) führten zudem ihre *Erfahrungen* (BF-SR-IF-EF) als beeinflussend an (Tabelle 123). Damit entfiel der größte Anteil der benannten Bewertungsfaktoren auf selbstreflexiver Ebene auf *interne Faktoren* (BF-SR-IF), von denen pro Proband*in im Mittel 2,9 verschiedene Aspekte benannt wurden.

Von den *externen Faktoren* (BF-SR-EF) wurden nur die *soziale Norm* (BF-SR-EF-SN) (33,3 %, 4 von 12) und die *Medien* (BF-SR-EF-ME) als beeinflussend wahrgenommen, was einen Durchschnitt von 0,6 verschiedenen reflektierten externen Faktoren pro Proband*in ergab.

Tabelle 123: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *selbstreflexive Faktoren* (BF-SR), *N* = 12

Kat.	1.UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
BF: Bewertungsfaktoren	SR: Selbstreflexive Faktoren	SR-IF: Interne Faktoren	BF-SR-IF-FW: Fachwissen	12	12	17	18
			BF-SR-IF-PR: Persönliche Relevanz	10	7	17	10
			BF-SR-IF-MW: Moral und Werte	1	2	1	3
			BF-SR-IF-MW: Handlungen	0	0	0	0
			BF-SR-IF-EM: Emotionen	2	5	2	7
			BF-SR-IF-EF: Erfahrungen	5	7	6	10
			BF-SR-IF-IN: Interesse	3	6	3	7
			BF-SR-IF-BV: Befangenheit, Vorurteile	2	4	2	4
			BF-SR-IF-PK: Persönlichkeit	0	0	0	0
			BF-SR-IF-KV: Kognitive Verzerrungen	0	0	0	0
			Summe SR-IF			**48**	**59**
			Einzelzählung SR-IF	**35**	**43**		
		SR-EF: Externe Faktoren	BF-SR-EF-SN: Soziale Norm	4	8	7	13
			BF-SR-EF-ME: Medien	3	3	3	3
			BF-SR-EF-BI: Studium/Bildung	0	3	0	4
			BF-SR-EF-PO: Politische Position	0	1	0	2
			BF-SR-EF-RL: Religion	0	0	0	0
			Summe SR-EF			**10**	**22**
			Einzelzählung SR-EF	**7**	**15**		
			Summe SR Gesamt			**58**	**81**
			Einzelzählung SR Gesamt	**42**	**58**		

In der Post-Erhebung stieg die Anzahl der unterschiedlichen berücksichtigten *internen Faktoren* (BF-SR-IF) von 2,9 auf durchschnittlich 3,6. Der größte Anteil fiel hier abermals auf das *Fachwissen* (BF-SR-IF-FW), welches alle Teilnehmenden als beeinflussend benannten. In den Kategorien *Befangenheit* (BF-SR-IF-BV) (33,3 %, 4 von 12) und *Interesse* (BF-SR-IF-IN) (50,0 %, 6 von 12) verdoppelte sich jeweils die Menge der Personen, die diese Faktoren als prägend anführten. Auch die Aspekte *Emotionen* (BF-SR-IF-EM) (42,7 %, 5 von 12) und *Erfahrungen* (BF-SR-IF-EF) (58,3 %, 7 von 12) traten in den Antworten vermehrt auf. Lediglich in der Kategorie

der *persönlichen Relevanz* (BF-SR-IF-PR) war eine Reduktion der Benennungen zu verzeichnen, was auf eine veränderte und weniger selbstfokussierte Schwerpunktsetzung zurückzuführen sein könnte.

Aufseiten der *externen Faktoren* (BF-SR-EF) war eine Erhöhung des Mittelwertes um mehr als das Doppelte erkennbar. Dieser stieg auf 1,25 und die Anzahl an Personen, die externe Faktoren als beeinflussend angaben, verdoppelte sich von 5 auf 10 (von 41,7 % auf 83,3 %). Die *soziale Norm* (BF-SR-EF-SN) wurde im Post-Test von 66,7 % (8 von 12) der Teilnehmenden als beeinflussend benannt. Jeweils drei Personen (25 %) führten *Medien* (BF-SR-EF-ME) und *Bildung* (BF-SR-EF-BI) als beeinflussende Aspekte an. Die *politische Position* (BF-SR-EF-PO) wurde von einer Person erwähnt.

Die Untersuchung der *epistemologischen Aspekte* (BF-EP) zeigte im Pre-Test im Schnitt 2,0 differente Aspekte pro Person. Jeweils 41,7 % (5 von 12) Teilnehmende gaben an, dass bei der Bewertung von Quellen vor allem die *Expertise einer Quelle* (BF-EP-EX), potenzielle *Interessenkonflikte* (BF-EP-IK) und die *Herkunft einer Information* (BF-EP-HK) eine vornehmliche Rolle spielen. Zusätzlich wurden die *Transparenz* (BF-EP-TR) von 33,3 % (4 von 12) und die *Übereinstimmung mit weiteren Quellen* (BF-EP-KO) von 25 % (3 von 12) der Befragten als relevante Faktoren gewichtet (Tabelle 124).

Die Bewertungsperspektiven im Umgang mit Quellen stellten sich im Post-Interview bei fast allen Teilnehmenden erweitert dar. Im Mittel führten diese nun 3,9 verschiedene Kriterien zur Bewertung von Quellen an. Alle Kategorien wurden in den Post-Interviews häufiger benannt, wobei die meisten Nennungen auf die Kategorien *Expertise* (BF-EP-EX) (75 %, 9 von 12), *Interessenkonflikte* (BF-EP-IK) und *Herkunft von Informationen* (BF-EP-HK) fielen. Die Hälfte der Befragten berücksichtigte zudem die Art, wie *Informationen präsentiert* (BF-EP-PR) werden, und inwiefern diese mit weiteren *Quellen übereinstimmen* (BF-EP-KO). Auch die *Aktualität* (BF-EP-AK) wurde vermehrt (33,3 %, 4 von 12) als wichtiges Kriterium angeführt.

Tabelle 124: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *epistemologische Faktoren* (BF-EP), $N = 12$

Kat.	1.UK 2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
BF: Bewertungsfaktoren	EP: Epistemologische Faktoren	BF-EP-EX: Autoritäten/Expertise	5	9	8	13
		BF-EP-AK: Aktualität	1	4	1	6
		BF-EP-IK: Interessenskonflikte	5	8	6	11
		BF-EP-PR: Präsentation	1	6	4	10
		BF-EP-HK: Herkunft der Information	5	7	7	9
		BF-EP-ME: Methodik	0	1	0	1
		BF-EP-KO: Konsistenzprüfung	3	6	5	6
		BF-EP-TR: Transparenz	4	5	5	5
		BF-EF-SE: Soziokulturelle Eingebundenheit	0	1	0	1
		Summe EP			**36**	**62**
		Einzelzählung EP	**24**	**47**		

Mit Blick auf die *naturwissenschaftlichen Aspekte* (BF-WI) zeigten sich im Pre-Test 3,2 verschiedene Perspektiven pro Proband*in. Besonders wurde die *Nachhaltigkeit* (BF-WI-NA) von 83,3 % (10 von 12) und die *Umsetzbarkeit* (BF-WI-UM) der Technologien von 66,7 % (8 von 12) der Teilnehmenden thematisiert. Die Kategorien *Effizienz* (BF-WI-EF), *Funktionsweise* (BF-WI-FT) und *Rohstoffherkunft* (BF-WI-RH) wurden von je 33,3 % (4 von 12) der teilnehmenden Personen zur Bewertung herangezogen (Tabelle 125).

Die bereits in der Pre-Erhebung bestehende mehrperspektivische Bewertungsgrundlage zeigte sich in der Post-Erhebung weiter ausdifferenziert und erweitert. Hier stieg die Anzahl differenter Bewertungsperspektiven im Mittel auf 4,25. Alle befragten Personen zogen den *Nachhaltigkeitsaspekt* (BF-WI-NA) als Bewertungsgrundlage hinzu. Je 83,3 % (10 von 12) berücksichtigten zudem die *Effizienz* (BF-WI-EF) und das *Kosten-Nutzen-Verhältnis* (BF-WI-KN). Auch die *Rohstoffherkunft* (BF-WI-RH) trat im Vergleich zum Pre-Test deutlich gehäufter auf (66,7 %, 8 von 12).

Tabelle 125: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Bewertungsfaktoren, *naturwissenschaftliche Faktoren* (BF-NW), *N* = 12

Kat.	1.UK	2. UK	3. UK	Summe Einzel Pre	Summe Einzel Post	Summe Mehrf. Pre	Summe Mehrf. Post
BF: Bewertungsfaktoren	NW: Naturwissenschaftliche Faktoren		BF-NW-WS: Wissenschaftliche Standards	1	0	1	0
			BF-NW-EF: Effizenz	4	10	5	20
			BF-NW-KA: Kostenabwägung	7	10	13	21
			BF-NW-NA: Auswirkungen & Nachhaltigkeit	10	12	27	29
			BF-NW-FT: Funktionsweise	4	2	5	3
			BF-NW-UM: Umsetzbarkeit	8	9	17	23
			BF-NW-RH: Rohstoffherkunft	4	8	6	12
			BF-NW-NO: Notwendigkeit	0	0	0	0
	Summe NW					**74**	**108**
	Einzelzählung NW			**38**	**51**		

Insgesamt erwiesen sich die Ergebnisse der offenen Fragestellungen, insbesondere jene zu den Bewertungsfaktoren, nach der Kontextualisierung als ausgesprochen ergiebig. Im Vergleich zu den nicht kontextualisierten Fragestellungen aus S3.F1&2 zeigte sich eine deutlich höhere inhaltliche Tiefe und Differenzierung in den Antworten. Gleichzeitig ließen sich auf qualitativer Ebene deutliche Unterschiede zwischen den Ergebnissen der Pre- und Posterhebung feststellen. Diese Differenzen wurden darüber hinaus auch quantitativ mithilfe des Wilcoxon-Tests auf Signifikanz überprüft (Tabelle 126).

In allen Bereichen gab es bei einem Signifikanzniveau von $\alpha = 0{,}05$ signifikante Unterschiede zwischen dem Pre- und Post-Test. Sowohl das *Verständnis der Selbstreflexion* (VS) ($z = -2{,}739$, $p = 0{,}004$, $r = 0{,}559$) als auch des *Kritischen Denkens* (VK) ($z = -3{,}228$, $p = 0{,}000$, $r = 0{,}659$) verbesserte sich im Post-Test signifikant und wies hohe Effektstärken auf. Genauso ließ das *Verständnis kognitiver Verzerrungen* (KV) eine signifikante Verbesserung erkennen ($z = -2{,}310$, $p = 0{,}031$, $r = 0{,}472$), die eine mittlere Effektstärke erreichte.

Tabelle 126: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der offenen Fragestellungen, $N = 12$

S5.F1&2 - Wilcoxon-Test

Teststatistikena	VS_Post - VS_Pre	VK_Post - VK_Pre	BFWI_Post - BFWi_Pre	BFEP_Post - BFEP_Pre	BFSR_Post - BFSR_Pre	BFSRIF_Post - BFSRIF_Pre	BFSREF_Post - BFSREF_Pre	KV_Post - KV_Pre
Z	-2,739b	-3,228b	-2,209b	-3,119b	-2,970b	-2,165b	-2,324b	-2,310b
Asymp. Sig. (2-seitig)	0,006	0,001	0,027	0,002	0,003	0,030	0,020	0,021
Exakte Sig. (2-seitig)	0,004	0,000	0,039	0,001	0,002	0,047	0,033	0,031
Exakte Sig. (1-seitig)	0,002	0,000	0,020	0,000	0,001	0,023	0,017	0,016
Punkt-Wahrscheinlichkeit	0,002	0,000	0,016	0,000	0,001	0,020	0,014	0,014
a Wilcoxon-Test								
b Die Summe der negativen Ränge ist gleich der Summe der positiven Ränge.								
c Basiert auf negativen Rängen.								
Effektstärke	0,559	0,659	0,451	0,637	0,606	0,442	0,474	0,472

Ebenso waren bei allen drei *Bewertungsfaktoren* (BF) statistisch signifikante Perspektiverweiterungen zu verzeichnen. Der Bereich der *wissensbezogenen Aspekte* (BF-WI) erreichte mit einer Teststatistik von $z = -2{,}209$ und einer exakten Signifikanz von $p = 0{,}039$ eine mittlere Effektstärke von $r = 0{,}451$. *Epistemologische Aspekte* (BF-EP) ($z = -3{,}119$, $p = 0{,}001$, $r = 0{,}637$) und *selbstreflexive Aspekte* (BF-SR) ($z = -2{,}970$, $p = 0{,}002$, $r = 0{,}606$) erzielten indes deutlich hohe Effektstärken. Die Differenzierung der beiden Subkategorien der Selbstreflexion erreichte sowohl signifikante Verbesserungen bei den *internen Faktoren* (BF-SR-IF) ($z = -2{,}165$, $p = 0{,}047$, $r = 0{,}442$) als auch den *externen Faktoren* (BF-SR-EF) ($z = -2{,}324$, $p = 0{,}033$, $r = 0{,}474$) bei mittlerer Effektstärke.

Dies untermauert die bereits qualitativ gewonnenen Eindrücke und die Annahme, dass die Kontextualisierung maßgeblich zur Aussagekraft des Instruments beigetragen hat.

Die Auswertung der geschlossenen Fragen zeigte im Fragenblock zur *Introspektion* (IS) des Pre-Tests einen Mittelwert von 3,083 Skalenpunkten über alle fünf Items hinweg (Tabelle 127). Die Mittelwerte der einzelnen Items waren zudem sehr vergleichbar und differierten höchstens um 0,416 Skalenpunkte. Die höchste Zustimmung erhielt die Auseinandersetzung mit den eigenen *Moral- und Wertvorstellungen* (IS2), während die niedrigste Zustimmung das Bewusstsein eigener *Gedanken und Gefühle* (IS1) erkennen ließ. Die Varianzen der einzelnen Items zeugten von geringen Abweichungen vom Mittelwert und damit einer gewissen Homogenität der Antworten.

Im Post-Test stieg der Mittelwert des Itemblocks auf 3,867 Skalenpunkte. Auch hier zeigte sich eine höhere Sensibilität für *Moral- und Wertvorstellungen* (IS2) im Gegensatz zu *Gedanken und Gefühlen* (IS1). Dennoch waren alle Mittelwerte der einzelnen Items um mindestens 0,5 Punkte erhöht. Die Minima der einzelnen Items stiegen ebenfalls von 1 bis 2 Skalenpunkten im Pre-Test auf 3 Skalenpunkte im Post-Test. Die Streuung der Mittelwerte lag auch im Post-Test unter 1,0, was einer geringen Abweichung entspricht. Diese Ergebnisse sprechen für eine Steigerung der subjektiv empfundenen Sensibilität für Einflussfaktoren und der damit verbundenen Reflexion.

Tabelle 127: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Introspektion* (IS), *N* = 12

Pre							
Code	**IS1**	**IS2**	**IS3**	**IS4**	**IS5**	**Summe IS**	**Summe rel. %**
Summe	35	40	36	38	36	185	740,0
Mittelwert	**2,917**	**3,333**	**3,000**	**3,167**	**3,000**	15,417	61,7
Stabw.	0,900	0,778	0,953	0,718	0,739	2,778	11,1
Varianz	0,811	0,606	0,909	0,515	0,545	7,720	123,5
Modus	2	4	3	3	3	18,0	72,0
Median	3	3,5	3	3	3	15,5	62,0
Min	2	2	1	2	2	10,0	40,0
Max	4	4	4	4	4	19,0	76,0

Post							
Code	**IS1**	**IS2**	**IS3**	**IS4**	**IS5**	**Summe IS**	**Summe rel. %**
Summe	44	49	47	45	47	232	928,0
Mittelwert	**3,667**	**4,083**	**3,917**	**3,750**	**3,917**	19,333	77,3
Stabw.	0,651	0,515	0,515	0,622	0,515	1,969	7,9
Varianz	0,424	0,265	0,265	0,386	0,265	3,879	62,1
Modus	4	4	4	4	4	18,0	72,0
Median	4	4	4	4	4	19,0	76,0
Min	3	3	3	3	3	16,0	64,0
Max	5	5	5	5	5	24,0	96,0

IS1:	Wie oft machen Sie sich Ihre eigenen Gedanken und Gefühle bewusst, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?
IS2:	Inwieweit setzen Sie sich mit deinen eigenen Moralvorstellungen und Werte auseinander, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?
IS3:	Wie häufig denken Sie über Ihre persönlichen Erfahrungen und Hintergründe nach, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?
IS4:	Inwiefern reflektieren Sie Ihre Interessen und Motivationen, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?
IS5:	Wie häufig hinterfragen Sie Ihre eigenen Vorurteile und Annahmen, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?

Tabelle 128: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Blocks *Introspektion* (IS), *N* = 12

Frage	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
IS1	0,031	-2,310b	0,472	Signifikant
IS2	0,008	-2,714b	0,554	Signifikant
IS3	0,008	-2,598b	0,530	Signifikant
IS4	0,063	-2,111b	0,431	Nicht signifikant
IS5	0,016	-2,495b	0,509	Signifikant

Die Überprüfung der Ergebnisse mittels des Wilcoxon-Tests machte signifikante Veränderungen in vier der fünf Items des Blocks *Introspektion* (IS) sichtbar (Tabelle 128). Nur das Item, welches die Reflexion *von Interessen und Motivationen* (IS4) erfragte, erreichte mit einer Teststatistik von z = -2,111 und einer exakten Signifikanz von p = 0,063 nicht das Signifikanzniveau von α = 0,05. Die anderen Items wiesen signifikante Veränderungen bei teilweise hohen Effektstärken auf.

Die subjektiv empfundene Bewertungsfähigkeit, welche durch das Item EF2 erfragt wurde, erreichte im Pre-Test einen Mittelwert von 3,083 Skalenpunkten, was auf eine moderate Selbsteinschätzung der rein fachlichen Bewertungsfähigkeit hindeutet (Tabelle 129).

Tabelle 129: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Einflussfaktoren 2 (EF2), *N* = 12

Code	EF2 Pre	EF2 Post	EF2 rel.Pre	EF2 rel.Post
Summe	37	47	740,0	940,0
Mittelwert	**3,083**	**3,917**	61,7	78,3
Stabw.	0,793	0,669	15,9	13,4
Varianz	0,629	0,447	251,5	178,8
Modus	3	4	60	80
Median	3	4	60	80
Min	2	3	40	60
Max	4	5	80	100

EF2:	Inwieweit glauben Sie, dass Sie in der Lage sind, vornehmlich auf fachlicher Ebene zu urteilen?

Diese Selbsteinschätzung stellte sich nach der Vermittlungseinheit im Post-Test optimistischer dar, was sich in einer Mittelwerterhöhung auf 3,917 ausdrückte. Die Streumaße ließen eine geringe Streuung der Antworten der Befragten erkennen. Dies legt nahe, dass die subjektive Einschätzung der eigenen Fähigkeit zur rein fachlichen Bewertung durch die explizite Thematisierung möglicher Einflussfaktoren sowie deren reflektierten Umgang positiv beeinflusst wird.

Tabelle 130: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Items EF2, *N* = 12

Frage	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
EF2	0,043	-2,153b	0,439	Signifikant

Der Wilcoxon-Test (Tabelle 130) stellte diese Veränderung ebenfalls als signifikant heraus und ermittelte eine mittlere Effektstärke von $r = 0{,}439$.

Die Untersuchung der subjektiv empfundenen *Einflussfaktoren* (EF4) und deren Reflexion zeigte im Pre-Test deutliche Unterschiede in den Gewichtungen der verschiedenen personenbezogenen Einflüsse auf (Tabelle 131). Während das *eigene Wissen* (EF4_12), die *eigenen Erfahrungen* (EF4_15) oder die *eigene Bildung* (EF4_5), also primär Faktoren, die sich auf die eigene Wissensgrundlage zur Bewertung beziehen, verstärkt als reflektiert angegeben wurden, wurden Faktoren, wie die *eigene Religion* (EF4_9) oder die *eigene Ethnie* (EF4_10), deutlich weniger in meinungsbildenden Prozessen berücksichtigt. Bei den Faktoren *eigenes Alter* (EF4_6) und *eigene soziale Schicht* (EF4_8) war zudem eine deutliche Streuung der Ergebnisse zu verzeichnen, was auf eine Heterogenität der Antworten in diesen Items hindeutet.

Im Post-Test waren fast alle Items hinsichtlich ihres Mittelwertes erhöht. Lediglich das Item EF4_6, welches die Reflexion des *eigenen Alters* erfragte, blieb unverändert. Die Reflexion der *eigenen Kultur* (EF4_11) und der *eigenen sozialen Schicht* (EF4_8) ließ die größten Veränderungen zwischen Pre- und Post-Test erkennen, aber auch die Items zur *politischen Einstellung* (EF4_14) und zur *Stimmung und Emotion* (EF4_2) zeigten Mittelwertunterschiede von mehr als 0,8. Die Streumaße der Items waren im Post-Test zudem deutlich reduziert, was auf eine höhere Homogenität der Antworten schließen lässt.

Tabelle 131: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Einflussfaktoren 4* (EF4), personenbezogene Einflüsse, $N = 12$

Pre																
Code	**EF4 1**	**EF4 2**	**EF4 3**	**EF4 4**	**EF4 5**	**EF4 6**	**EF4 7**	**EF4 8**	**EF4 9**	**EF4 10**	**EF4 11**	**EF4 12**	**EF4 13**	**EF4 14**	**EF4 15**	**EF4 16**
Summe	35	35	41	41	46	31	22	29	13	19	22	49	44	38	46	33
MW	**2,917**	**2,917**	**3,417**	**3,417**	**3,833**	**2,583**	**1,833**	**2,417**	**1,083**	**1,583**	**1,833**	**4,083**	**3,667**	**3,167**	**3,833**	**2,750**
Stabw.	1,240	0,900	0,900	0,793	0,718	1,379	1,030	1,443	0,289	0,669	0,577	0,793	0,651	0,835	0,718	0,965
Varianz	1,538	0,811	0,811	0,629	0,515	1,902	1,061	2,083	0,083	0,447	0,333	0,629	0,424	0,697	0,515	0,932
Modus	3	3	4	3	4	2	1	1	1	1	2	4	4	3	4	2
Median	3	3	3,5	3	4	2	1,5	2	1	1,5	2	4	4	3	4	3
Min	1	2	2	2	3	1	1	1	1	1	1	3	3	2	3	1
Max	5	5	5	5	5	5	4	5	2	3	3	5	5	5	5	4

Post																
Code	**EF4 1**	**EF4 2**	**EF4 3**	**EF4 4**	**EF4 5**	**EF4 6**	**EF4 7**	**EF4 8**	**EF4 9**	**EF4 10**	**EF4 11**	**EF4 12**	**EF4 13**	**EF4 14**	**EF4 15**	**EF4 16**
Summe	43	45	50	49	52	31	30	40	20	24	35	55	56	48	53	42
MW	**3,583**	**3,750**	**4,167**	**4,083**	**4,333**	**2,583**	**2,500**	**3,333**	**1,667**	**2,000**	**2,917**	**4,583**	**4,667**	**4,000**	**4,417**	**3,500**
Stabw.	0,793	0,754	0,389	0,669	0,492	0,669	1,000	0,888	0,778	0,739	0,793	0,669	0,492	0,739	0,515	1,000
Varianz	0,629	0,568	0,152	0,447	0,242	0,447	1,000	0,788	0,606	0,545	0,629	0,447	0,242	0,545	0,265	1,000
Modus	4	3	4	4	4	2	3	3	1	2	3	5	5	4	4	4
Median	4	4	4	4	4	2,5	2,5	3	1,5	2	3	5	5	4	4	3,5
Min	2	3	4	3	4	2	1	2	1	1	2	3	4	3	4	2
Max	5	5	5	5	5	4	4	5	3	3	4	5	5	5	5	5

EF4:	Wie sehr hinterfragen Sie folgende Faktoren, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	
Personenbezogene Einflüsse	EF4_1	Eigene Persönlichkeit
	EF4_2	Stimmung und Emotion
	EF4_3	Werte und Moralvorstellungen
	EF4_4	Eigene Grundeinstellung
	EF4_5	Eigene Bildung
	EF4_6	Eigenes Alter
	EF4_7	Eigenes Geschlecht
	EF4_8	Eigene soziale Schicht
	EF4_9	Eigene Religion
	EF4_10	Eigene Ethnie
	EF4_11	Eigene Kultur
	EF4_12	Wissen
	EF4_13	Medialer Kontext
	EF4_14	Politische Einstellung
	EF4_15	Erfahrungen und Erinnerungen
	EF4_16	Selbstkonzept und Selbstwertgefühl

Die Items mit Bezug zu kognitiven Verzerrungen ließen im Pre-Test ebenfalls Schwerpunkte bei den reflektierten Faktoren erkennen (Tabelle 132). So wurden beispielsweise *Gruppendynamiken* (EF4_21) und *Übergeneralisierungen* (EF4_25) weniger reflektiert, wohingegen die *Verfügbarkeit von Alternativen* (EF4_19) und die *Motivation und Bedürfnisse* (EF4_18) als häufiger reflektiert angegeben wurden.

Im Post-Test wiederum war die Ausprägung der Antworten auf die Items soziale *Normen und Erwartungen* (EF4_17) und *Übergeneralisierung* (EF4_25) stark gesteigert, mit Mittelwertunterschieden von über 1,0 Skalenpunkten. Auch *Gruppendynamiken* (EF4_21), *Framing und Präsentation* (EF4_23), *selektive Wahrnehmung und Bestätigungsfehler* (EF4_24) sowie die *Komplexität der Information* (EF4_26) ließen Mittelwerterhöhungen von über 0,75 Skalenpunkten erkennen. Auch die Streuung der Ergebnisse blieb im niedrigen Bereich, mit Ausnahme des Items *Intuition und Bauchgefühl* (EF4_20), welches eine Varianz von 1,720 aufwies.

Tabelle 132: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Einflussfaktoren 4* (EF4), kognitive Verzerrungen, *N* = 12

Pre												
Code	**EF4 17**	**EF4 18**	**EF4 19**	**EF4 20**	**EF4 21**	**EF4 22**	**EF4 23**	**EF4 24**	**EF4 25**	**EF4 26**	**Summe EF4**	**Summe EF4 rel.**
Summe	35	41	42	39	31	34	38	36	31	40	911	700,8
MW	**2,917**	**3,417**	**3,500**	**3,250**	**2,583**	**2,833**	**3,167**	**3,000**	**2,583**	**3,333**	75,917	58,4
Stabw.	0,900	0,900	0,798	0,866	0,793	0,577	0,835	1,044	0,669	0,888	11,016	8,5
Varianz	0,811	0,811	0,636	0,750	0,629	0,333	0,697	1,091	0,447	0,788	121,356	71,8
Modus	2	4	4	3	2	3	4	2	2	3	#NV	#NV
Median	3	3,5	3,5	3	2	3	3	3	2,5	3	75,5	58,1
Min	2	2	2	2	2	2	2	2	2	2	61,0	46,9
Max	4	5	5	5	4	4	4	5	4	5	100,0	76,9

Post												
Code	**EF4 17**	**EF4 18**	**EF4 19**	**EF4 20**	**EF4 21**	**EF4 22**	**EF4 23**	**EF4 24**	**EF4 25**	**EF4 26**	**Summe EF4**	**Summe EF4 rel.**
Summe	48	45	46	43	41	42	49	45	43	49	1124	864,6
MW	**4,000**	**3,750**	**3,833**	**3,583**	**3,417**	**3,500**	**4,083**	**3,750**	**3,583**	**4,083**	93,667	72,1
Stabw.	0,739	0,866	0,718	1,311	0,793	0,674	0,669	0,754	0,900	0,900	6,513	5,0
Varianz	0,545	0,750	0,515	1,720	0,629	0,455	0,447	0,568	0,811	0,811	42,424	25,1
Modus	4	4	4	3	3	3	4	3	3	4	93,0	71,5
Median	4	4	4	3,5	3	3	4	4	3,5	4	93,0	71,5
Min	3	2	3	1	2	3	3	3	2	2	84,0	64,6
Max	5	5	5	5	5	5	5	5	5	5	105,0	80,8

EF4:	Wie sehr hinterfragen Sie folgende Faktoren, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	
Kognitive Verzerrungen	EF4_17	Soziale Normen und Erwartungen
	EF4_18	Motivation und Bedürfnisse
	EF4_19	Verfügbarkeit von Alternativen
	EF4_20	Intuition und Bauchgefühl
	EF4_21	Gruppendynamik und Konformitätsdruck
	EF4_22	Erwartungen und Vorurteile
	EF4_23	Framing und Präsentation von Information
	EF4_24	Selektive Wahrnehmung und Bestätigungsfehler
	EF4_25	Übergeneralisierung und Kategorisierung
	EF4_26	Komplexität der Information

Tabelle 133: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der Items zu EF4, $N = 12$

Frage	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
EF4_1	0,055	-2,126b	0,434	Nicht signifikant
EF4_2	0,018	-2,486b	0,507	Signifikant
EF4_3	0,047	-2,165b	0,442	Signifikant
EF4_4	0,035	-2,309b	0,471	Signifikant
EF4_5	0,070	-2,121b	0,433	Nicht signifikant
EF4_6	1,000	,000c	0,000	Nicht signifikant
EF4_7	0,031	-2,271b	0,464	Signifikant
EF4_8	0,031	-2,309b	0,471	Signifikant
EF4_9	0,031	-2,333b	0,476	Signifikant
EF4_10	0,125	-1,890b	0,386	Nicht signifikant
EF4_11	0,002	-2,919b	0,596	Signifikant
EF4_12	0,156	-1,730b	0,353	Nicht signifikant
EF4_13	0,002	-2,972b	0,607	Signifikant
EF4_14	0,039	-2,226b	0,454	Signifikant
EF4_15	0,039	-2,333b	0,476	Signifikant
EF4_16	0,016	-2,460b	0,502	Signifikant
EF4_17	0,004	-2,754b	0,562	Signifikant
EF4_18	0,219	-1,633b	0,333	Nicht signifikant
EF4_19	0,125	-2,000b	0,408	Nicht signifikant
EF4_20	0,344	-1,190b	0,243	Nicht signifikant
EF4_21	0,031	-2,308b	0,471	Signifikant
EF4_22	0,055	-2,126b	0,434	Nicht signifikant
EF4_23	0,008	-2,598b	0,530	Signifikant
EF4_24	0,047	-2,165b	0,442	Signifikant
EF4_25	0,004	-2,762b	0,564	Signifikant
EF4_26	0,045	-2,179b	0,445	Signifikant

Die Überprüfung auf signifikante Unterschiede zwischen Pre- und Post-Test zeigte beim Großteil der Items von EF4 (65 %, 17 von 26) signifikante Veränderungen in positive Richtung (Tabelle 133). Unter den Aspekten, die unverändert blieben, waren unter anderem die Reflexion des *eigenen Alters* (EF4_6), der *eigenen Ethnie* (EF4_10) oder des *eigenen Bauchgefühls* (EF4_20). Diese Aspekte bildeten zum Teil sehr komplexe Konstrukte, wie die *eigene Persönlichkeit* (EF4_1), ab oder lagen bereits im Pre-Test erhöht vor, wie die Reflexion des *eigenen Wissens* (EF4_12) oder die *Verfügbarkeit von Alternativen* (EF4_19). Die höchsten Effektstärken waren hingegen bei den Items *medialer Kontext* (EF4_13), *Übergeneralisierung und Kategorisierung* (EF4_25), *eigene Kultur* (EF4_11) und *soziale Normen und Erwartungen* (EF4_ 17) zu verzeichnen, welche alle Signifikanzen von unter 0,01 erreichten.

Diese Einflussfaktoren wurden im Pre-Test im Fragenblock *Bewertung und Gewichtung* (BG) von den Teilnehmenden noch als wenig beeinflussend wahrgenommen und infolgedessen auch mäßig berücksichtigt (Tabelle 134). Erkennbar wurde dies in den niedrigen Mittelwerten beider Items (BG1 und BG2) von 2,250 und 2,667 Skalenpunkten.

Tabelle 134: S5.F1&2 - Ergebnisse der Fragebogenerhebung, *Bewertung und Gewichtung* (BG), *N* = 12

Pre				
Code	**BG1**	**BG2**	**Summe BG**	**Summe BG rel. %**
Summe	27	32	59	590,0
Mittelwert	**2,250**	**2,667**	4,917	49,2
Stabw.	0,622	1,073	1,443	14,4
Varianz	0,386	1,152	2,083	208,3
Modus	2	3	6,0	60,0
Median	2	3	5,0	50,0
Min	1	1	2,0	20,0
Max	3	4	7,0	70,0

Post				
Code	**BG1**	**BG2**	**Summe BG**	**Summe BG rel. %**
Summe	40	43	83	830,0
Mittelwert	**3,333**	**3,583**	6,917	69,2
Stabw.	0,651	0,669	0,900	9,0
Varianz	0,424	0,447	0,811	81,1
Modus	3	3	6,0	60,0
Median	3	3,5	7,0	70,0
Min	2	3	6,0	60,0
Max	4	5	8,0	80,0

BG1:	Wie stark schätzen sie den Einfluss der vorhin genannten Faktoren (Bauchgefühl, Stimmung usw.) im Vergleich zum Fachwissen, wenn Sie ein Thema mit naturwissenschaftlichem Bezug bewerten sollen?
BG2:	Inwieweit wägen Sie diese Faktoren (Bauchgefühl. Stimmung usw.) bewusst gegen Ihr Fachwissen ab, wenn sie ein Thema mit naturwissenschaftlichem Bezug bewerten sollen?

Diese Durchschnittswerte stiegen im Post-Test bei der *Einschätzung der Einflüsse* (BG1) von 2,250 auf 3,333 und bei der *Berücksichtigung dieser in Bewertungsprozessen* (BG2) von 2,667 auf 3,583. Auch die Minima und der Median stiegen im Vergleich zum Pre-Test an und die Varianz zeigte eine geringe Streuung der Antworten.

Die erkennbaren Mittelwertunterschiede erreichten bei der Überprüfung mithilfe des Wilcoxon-Tests deutlich das Signifikanzniveau und wiesen hohe Effektstärken auf (Tabelle 135). Dies lässt den Schluss zu, dass ein erhöhtes Bewusstsein für die Einflüsse verschiedener Faktoren erreicht werden konnte, welches sich subjektiv auch auf die Bewertungssituationen überträgt.

Tabelle 135: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test des Blocks *Bewertung und Gewichtung (BG)*, $N = 12$

Frage	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
BG1	0,002	-2,919b	0,596	Signifikant
BG2	0,016	-2,495b	0,509	Signifikant

Die *Self-Reflection and Insight Scale* (SRIS) zeigte im Pre-Test auf der Subskala der *Selbstreflexion* (SR) einen Mittelwert von 4,611 bei moderater bis hoher Streuung der Antworten (Tabelle 136). Der Mittelwert der Subskala der *Einsicht* (IN) lag nur knapp darunter bei 4,583 und wies ebenfalls eine moderate bis hohe Streuung der Antworten auf. Damit stellte sich die Selbsteinschätzung der Neigung, selbstreflexives Verhalten zu zeigen sowie daraus Einsichten gewinnen zu können, sehr heterogen mit leicht positiver Tendenz dar.

Tabelle 136: S5.F1&2 - Ergebnisse der Fragebogenerhebung, SRIS nach Silvia (2021), *N* = 12

Pre

Code	SR3	SR5	SR6	SR8	SR9	SR10	Summe SR	Summe SR rel.
Summe	48	56	56	55	58	59	332	790,5
Mittelwert	**4,000**	**4,667**	**4,667**	**4,583**	**4,833**	**4,917**	27,667	65,9
Stn. Abw.	1,279	1,826	1,155	1,564	1,528	1,084	6,358	15,1
Varianz	1,636	3,333	1,333	2,447	2,333	1,174	40,424	229,2
Modus	5	6	5	5	6	5	33,0	78,6
Median	4	5	5	5	5,5	5	29,5	70,2
Min	2	2	3	1	2	3	19,0	45,2
Max	6	7	6	6	6	7	36,0	85,7

Post

Code	SR3	SR5	SR6	SR8	SR9	SR10	Summe SR	Summe SR rel.
Summe	58	69	70	69	75	72	413	983,3
Mittelwert	**4,833**	**5,750**	**5,833**	**5,750**	**6,250**	**6,000**	34,417	81,9
Stn. Abw.	1,467	1,215	1,115	0,965	0,866	1,414	5,143	12,2
Varianz	2,152	1,477	1,242	0,932	0,750	2,000	26,447	149,9
Modus	5	7	6	6	7	7	36,0	85,7
Median	5	5,5	6	6	6,5	7	35,5	84,5
Min	3	4	4	4	5	3	23,0	54,8
Max	7	7	7	7	7	7	42,0	100,0

Pre

Code	IN2	IN4	IN5	IN6	IN7	IN8	Summe IN	Summe IN rel.	Summe SR+IN Gesamt	Summe SR+IN Ges. rel.
Summe	60	55	56	52	56	51	330	785,7	662	788,1
Mittelwert	**5,000**	**4,583**	**4,667**	**4,333**	**4,667**	**4,250**	27,500	65,5	55,167	65,7
Stn. Abw.	1,128	1,165	1,155	1,826	1,775	1,815	6,895	16,4	10,321	12,3
Varianz	1,273	1,356	1,333	3,333	3,152	3,295	47,545	269,5	106,515	151,0
Modus	5	4	5	3	6	2	29,0	69,0	54,0	64,3
Median	5	4,5	5	4	5	4,5	29,0	69,0	53,0	63,1
Min	2	3	2	1	2	2	15,0	35,7	37,0	44,0
Max	6	7	6	7	7	7	38,0	90,5	70,0	83,3

Post

Code	IN2	IN4	IN5	IN6	IN7	IN8	Summe IN	Summe IN rel.	Summe Gesamt	Summe Ges. rel.
Summe	69	67	63	63	63	62	387	921,4	800	952,4
Mittelwert	**5,750**	**5,583**	**5,250**	**5,250**	**5,250**	**5,167**	32,250	76,8	66,667	79,4
Stn. Abw.	0,866	0,669	1,055	1,712	1,357	1,267	5,562	13,2	9,500	11,3
Varianz	0,750	0,447	1,114	2,932	1,841	1,606	30,932	175,4	90,242	127,9
Modus	6	6	6	6	6	4	38,0	90,5	62,0	73,8
Median	6	6	6	6	5,5	5	34,0	81,0	69,0	82,1
Min	4	4	3	2	3	3	22,0	52,4	46,0	54,8
Max	7	6	6	7	7	7	38,0	90,5	80,0	95,2

Pre	Mittelwert SR Pre	Mittelwert IN Pre	Mittelwert Gesamt Pre
	4,611	4,583	4,597
Post	**Mittelwert SR Post**	**Mittelwert IN Post**	**Mittelwert Gesamt Post**
	5,736	5,375	5,556

Die Selbsteinschätzung, *selbstreflexives Verhalten* zu zeigen (SR), stieg im Post-Test im Mittel um 1,125 Skalenpunkte auf 5,736 an und auch die Subskala der *Einsicht* (IN) stieg um 0,729 Punkte auf 5,375. Daraus lässt sich eine subjektiv gesteigerte Tendenz, selbstreflexives Verhalten zu zeigen und daraus Einsichten gewinnen zu können, ableiten.

Die Überprüfung der Ergebnisse mittels des Wilcoxon-Tests zeigte bei nahezu allen Items der SRIS signifikante Veränderungen in positive Richtung mit moderaten bis hohen Effektstärken (Tabelle 137). Lediglich das Item IN5 wies keine signifikanten Veränderungen auf. Da jedoch beim SRIS der Fokus auf der Gesamtbetrachtung der einzelnen Subskalen liegt, kann durchaus festgehalten werden, dass auch auf Ebene der *Einsicht* (IN) eine deutliche Veränderung zu verzeichnen war.

Tabelle 137: S5.F1&2 - Ergebnisse der Fragebogenerhebung, Wilcoxon-Test der SRIS, $N = 12$

Frage	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
SR3	0,008	-2,640b	0,539	Signifikant
SR5	0,014	-2,511b	0,513	Signifikant
SR6	0,002	-2,889b	0,590	Signifikant
SR8	0,002	-2,913b	0,595	Signifikant
SR9	0,002	-2,850b	0,582	Signifikant
SR10	0,016	-2,414b	0,493	Signifikant
IN2	0,033	-2,324b	0,474	Signifikant
IN4	0,016	-2,489b	0,508	Signifikant
IN5	0,148	-1,732b	0,354	Nicht signifikant
IN6	0,031	-2,232b	0,456	Signifikant
IN7	0,031	-2,333b	0,476	Signifikant
IN8	0,028	-2,299b	0,469	Signifikant

Analog zu Studie S3.F1&2 wurde der CRT auch hier nur im Pre-Test durchgeführt, da die Kenntnis der Fragen eine erneute Durchführung verfälschen würde. Zudem konnten hier $N = 16$ Datensätze ausgewertet werden, da an der Pre-Erhebung noch 16 Personen beteiligt waren. Vier dieser Teilnehmenden nahmen jedoch nicht an der

Post-Erhebung teil, weshalb diese nicht im Pre-Post-Vergleich der anderen Itemblöcke auftauchten.

Die Ergebnisse des *Cognitive Reflection Tests* (CRT) zeigten ein vergleichbares Muster zu den Ergebnissen der anderen Erhebungen und der Referenzwerte von Frederick (2005) (vgl. Kap. 4.2.3, Tabelle 9) (Tabelle 138). Der Mittelwert lag bei 1,563 korrekten Antworten pro Person, wobei 18,8 % (3 von 16) der Teilnehmenden alle Fragen korrekt beantworten konnten und 12,5 % (2 von 16) keine der Fragen korrekt lösten. Der größte Teil der Befragten verteilte sich somit im mittleren Bereich von einer (37,5 %, 6 von 16) oder zwei (31,3 %, 5 von 16) korrekten Antworten. Dies lässt die Annahme zu, dass bei 50 % der Befragten (keine oder eine korrekte Antwort) nur eine mäßige bis schlechte Erstimpulsinhibition zugunsten heuristischer Urteile vorlag.

Tabelle 138: S5.F1&2 - Ergebnisse der Fragebogenerhebung, CRT nach Frederick (2005), *N* = 16

Code	CRT1	CRT2	CRT3	Korrekt	%
Korrekt	5	12	8	25	**52,1**

Verteilung der Ergebisse in %			
12,5	37,5	31,3	18,8
0	1	2	3
Low			High

Resümierend zeugte die durchgeführte Fragebogenerhebung S5.F1&2 von signifikanten Veränderungen, sowohl im Hinblick auf das Verständnis zentraler Begrifflichkeiten und die Anzahl zur Bewertung hinzugezogener Aspekte als auch im subjektiven Empfinden der einzelnen Teilnehmenden bezüglich ihrer Tendenz zur Selbstreflexion und der Einbeziehung dieser in den Bewertungsprozess. Dabei ermöglichte die Kombination aus offenen und geschlossenen Fragen eine zur Untersuchung dieser komplexen Denk- und Bewertungsstrukturen notwendige Tiefe und Breite.

Die implementierten Kontextualisierungen der offenen Fragen zu den Bewertungsfaktoren zeigten eine deutliche Steigerung des Erkenntnisgewinns, insbesondere in den Bereichen der epistemologischen und wissensbezogenen Aspekte. Dies legt die Vermutung nahe, dass konkrete Anforderungssituationen eine zentrale Voraussetzung für die Untersuchung von Denk- und Bewertungsprozessen darstellen, da sie eine Annäherung an eine kompetenzorientierte Erhebung ermöglichen. Dennoch bleibt auch das offene und kontextualisierte Fragenformat in seiner Tiefe limitiert,

weshalb eine Triangulation mit einer ergänzenden Erhebungsform, wie dem Führen von Interviews, weiterhin notwendig bleibt.

5.5.3 Interkoderreliabilität

Auch in der Studie S5 wurde die Interkoderreliabilität der Kategorien der Kodierleitfäden überprüft, indem mindestens 50 % der Datensätze unabhängig von einem zweiten Interkoder mit Expertise in der Chemiedidaktik kodiert wurden.

Fünf der sieben Datensätze der Interviewstudie S5.I1&2, bestehend aus Pre- und Post-Test, wurden zufällig ausgewählt und dem Interkoder zur Verfügung gestellt. Die Kategorien zu den *Verständnissen* (SR-V und KD-V) sowie zu den *Bewertungsaspekten* (SR-A, EP-A, NW-A) wiesen sehr gute Übereinstimmungen der Kodierungen von Koder und Interkoder mit Cohens-Kappa-Werten von über $\varkappa = 0{,}800$ auf (Tabelle 139). Nur die Kategorie der *fachlichen Bewertung* (NW-FB) erreichte einen niedrigeren Wert von $\varkappa = 0{,}649$, was jedoch immer noch einer hohen Übereinstimmung entspricht und auf die geringe Anzahl an Kodierungen von Koder und Interkoder zurückzuführen sein kann.

Tabelle 139: S3.I1&2 – Interkoderreliabilität der Interviewerhebung

Kategorie	Cohens-Kappa ($\varkappa$)
Aspekte naturwissenschaftlicher Bewertung (NW-A)	$\varkappa = 0{,}887$
Aspekte der Selbstreflexion (SR-A)	$\varkappa = 0{,}895$
Aspekte epistemologischen Verständnisses (EP-A)	$\varkappa = 0{,}884$
Verständnis Selbstreflexion (SR-V)	$\varkappa = 0{,}894$
Verständnis Kritisches Denken (KD-V)	$\varkappa = 0{,}887$
Selbsteinschätzung Fachwissen (NW-SF)	$\varkappa = 0{,}846$
Fachliche Bewertung (NW-FB)	$\varkappa = 0{,}649$
Fähigkeit zur Bewertung (FB)	$\varkappa = 0{,}847$

Zur Verdeutlichung unterschiedlicher Bewertungen von Koder und Interkoder werden im Folgenden exemplarische Stellen angeführt, die in der Kodierkonferenz diskutiert wurden.

So wurde beispielsweise die Stelle „Aber deshalb würde ich sagen, bezogen auf mich selbst bin ich kritisch, auch im Umgang mit meinen eigenen Positionen und Mei-

nung“ (S5.I1-WS23-D-M01-AB30VT, Pos. 70) vom Interkoder bei der Begriffsdefinition *Kritischen Denkens* (KD-V) der Kategorie der *Selbstverbesserung* (KD-V-NA-SV) zugeordnet, da die Aussage den kritischen Umgang mit sich selbst thematisierte. Im Gegensatz dazu wies der Koder dieser Stelle die Kategorie *Nicht definierbar* (KD-V-NA-ND) zu, da Kritisches Denken mit dem Wort „kritisch“ selbst erklärt wurde und daraus kein Verständnis abgeleitet werden konnte. In der Kodierkonferenz wurde sich geeinigt, die Kodierung *Nicht definierbar* (KD-V-NA-ND) beizubehalten.

Ein weiteres Beispiel stellt die Aussage „[…] Reflexion in der Hinsicht, dass ich da nochmal meine Position überdenken muss […]“ (S5.I1-WS23-D-M01-AB30VT, Pos. 64) dar, welche vom Koder als *Introspektion* (SR-V-ET-IS) kodiert wurde, da das Überdenken der eigenen Position eine innere Einsichtnahme darstellt. Der Interkoder ordnete die Aussage hingegen der Kategorie *Hinterfragen* (SR-V-IT-HF) zu, da das Überdenken der eigenen Position nur vage formuliert ist. Dieser Position wurde sich in der Kodierkonferenz angeschlossen und die Passage als *Hinterfragen* (SR-V-IT-HF) kodiert.

Von der Fragebogenerhebung S5.F1&2 wurden sechs der zwölf Datensätze interkodiert und auch hier hohe Übereinstimmungen der Kategorien erreicht (Tabelle 140).

Tabelle 140: S3.F1&2 – Interkoderreliabilität der Fragebogenerhebung

Kategorie	Cohens-Kappa ($\varkappa$)
Bewertungsfaktoren (BF)	$\varkappa = 0{,}898$
Verständnis Selbstreflexion (VS)	$\varkappa = 0{,}874$
Verständnis Kritisches Denken (VK)	$\varkappa = 0{,}886$
Verständnis kognitive Verzerrungen (KV)	$\varkappa = 0{,}902$

Alle Kategorien erreichten Übereinstimmungen von über $\varkappa = 0{,}800$, was einer hohen Trennschärfe entspricht.

Zusammengefasst zeigten sich beide Instrumente, das Interview und der Fragebogen, hinsichtlich ihrer Kodierungen und Kategorien, als überaus trennscharf und reliabel, was die Verlässlichkeit der Kodierleitfäden und deren unabhängige Anwendbarkeit unterstreicht. Dabei stellte sich die an Mayring (2010) angelehnte Kodierkonferenz in der Ausschärfung der Kodierregeln als wichtiges Instrument zur Sicherstellung konsistenter Kodierungen bei komplexeren Formulierungen heraus und damit als sinnvolle Ergänzung zur eingesetzten qualitativen Inhaltsanalyse nach Kuckartz (2016) zugunsten einer höheren Intersubjektivität und Reliabilität.

5.5.4 Rückmeldungen der Teilnehmenden zur Vermittlungseinheit

Analog zur Studie S3 fand in der Studie S5 zur Überprüfung der Funktionalität des Instruments die in Kapitel 4.2.5 beschriebene Vermittlungseinheit statt. Trotz des geringeren Zeitumfangs (vgl. Kap. 5.5) konnte die Systemisch-reflexive Stellungnahme (SrS) mit den Teilnehmenden des Seminars durchgeführt werden. In diesem Rahmen hatten die Teilnehmenden die Möglichkeit, ihre Prozessreflexion zur vierten Aufgabe der SrS nach Abschluss der Sitzung einzureichen. Im Folgenden werden zwei exemplarische Rückmeldungen der Studierenden dargestellt, um einen Einblick in die Reflexionen der Teilnehmenden zu ermöglichen[15].

Eine Person reflektierte insbesondere die Rolle politischer Zugehörigkeit auf den eigenen Bewertungsprozess und wie Algorithmen von Suchmaschinen die Informationsaneignung unbewusst beeinflussen können:

> Während meiner Recherche bin ich auf viele verschiedene Artikel und wissenschaftliche Berichte gestoßen, die das Thema aus unterschiedlichen Blickwinkel betrachteten. Besonders auffällig waren die vielen Artikel, welche sich mit den Ansichten der politischen Parteien beschäftigt haben. Hier ist auch die Gefahr groß, dass sich die eigene Meinung zum Thema von der politischen Zuordnung beeinflussen lässt. Mir ist aufgefallen, dass ich zu diesem Thema besonders Artikel aufgezeigt bekommen habe, die sich mit Parteien oder Organisationen beschäftigen, welche meine politische Meinung widerspiegeln. Ohne mich also großartig mit dem Thema beschäftigt zu haben, konnte ich bereits sehen, wie „mein" politisches Lager zu dieser Thematik steht. So habe ich direkt eine Grundhaltung gegenüber diesem Thema entwickelt, ohne wirklich verstanden zu haben, wie diese Haltung zu begründen ist. Andere Meinungen aus anderen politischen Lagern wurden dann direkt negativer gewertet. Auch bei der Auseinandersetzung mit wissenschaftlichen Themen konnte ich sehen, wie ich vorzugsweise Artikel gesucht habe, die meine Meinung und meiner politischen Haltung entsprachen (S5-SrS-1).

Eine weitere Person thematisierte den durch die Vermittlungseinheit erzielten persönlichen Erkenntnisgewinn und reflektierte dessen Einfluss auf die eigenen Denk- und Bewertungsprozesse:

> Es war für mich ein Lernprozess, bewusst eine kritische Haltung einzunehmen und zu erkennen, dass die Auswahl und Bewertung der Quellen entscheidend ist. Es war nicht nur wichtig, die Fachlichkeit einer Quelle zu berücksichtigen, sondern auch zu erkennen, dass persönliche Perspektiven und Meinungen eine Rolle spielen. Die Herausforderung bestand darin, eine ausgewogene und fachlich fundierte Meinung zu finden, besonders wenn man

[15] Die Zitate wurden im Original belassen, wobei aus Gründen der Lesbarkeit auf die Kennzeichnung sprachlicher, orthografischer oder grammatischer Fehler verzichtet wurde.

bedenkt, dass Suchmaschinen wie Google oft von algorithmischen Vorlieben beeinflusst werden. Insgesamt habe ich durch diese Erfahrung nicht nur meine Fähigkeiten zur Recherche von fachlichen Quellen verbessert, sondern auch gelernt, wie wichtig es ist, bewusst mit nichtfachlichen Quellen umzugehen. Diese Erkenntnis hat nicht nur meine Herangehensweise an die aktuelle Aufgabe beeinflusst, sondern wird auch meine zukünftigen Forschungsansätze prägen. Es war eine persönliche Erfahrung des Lernens und der Anpassung, die meine Fähigkeiten im Umgang mit Informationsquellen vertieft hat (S5-SrS-2).

Auch in den Interviews fanden sich wiederholt Aussagen, die auf die Vermittlungseinheit und die explizite Thematisierung Kritischen Denkens Bezug nahmen und verdeutlichten, inwiefern diese das Denken der Teilnehmenden beeinflusst haben:

[W]ie wir die Partnerarbeit gemacht haben, wo wir diese Reflexionsfragen dazu hatten, zu dem Thema, zu dem wir uns recherchiert haben. Ich habe da noch mal ein paar Ebenen mehr entdeckt, die ich nicht so auf dem Schirm hatte, weil wie gesagt, zum Beispiel, ob die politische Einstellung da irgendwie mit reinspielt. Ist ja manchmal auch die Frage, was ja manche Menschen tun. Habe ich eine Einstellung und suche mir die Partei dazu oder kann die Partei mich auch zu einer Einstellung bringen? (S5.I2-WS23-D-M01-NB13SH, Pos. 56).

Ich fand es sehr wichtig oder etwas, was es was mir gebracht hat, durch diese Unterrichtsreihe, war, in einer Argumentation den eigenen Standpunkt oder bzw. den eigenen Hintergrund zu beleuchten, weil das ist etwas, was ich vorher noch nie gemacht habe und das habe ich neu dazugelernt (S5.I2-WS23-C-M01-SJ21BO, Pos. 76).

Also ich habe da immer so ein bisschen das Affektive, weil ich weiß, dass ich dazu neige, und das Intuitive so im Vordergrund, dass ich bei der Selbstreflexion manchmal denke: Bist du da nicht, neigst du nicht mal wieder dazu, hast du da nicht wieder vielleicht einen Fehler gemacht oder hast du da einen Fehler in der Denkweise? Dass ich mich immer sozusagen zu dem, zu dem naturwissenschaftlichen oder fachlichen Denken und dem rationaleren Aspekt der Kognition immer wieder hinbringen muss und bedenken muss: Ja, Kognition wird auch durch Affekte beeinflusst und man muss manchmal einen Schritt wieder zurückgehen und das Reflektieren und wieder herausnehmen, um eine wirklich rationale Analyse machen zu können (S5.3-WS23-E-M01-NB13SH, Pos. 32).

[…]

Kritisches Denken oder kritisch Sein hat für mich vorher auch immer eher die Komponente rational gehabt. Wobei mittlerweile […] kritisch Sein für mich auch dann wiederum bedeutet, dass da irgendwo die Selbstreflexion mit hineinspielt, dass ich nämlich einfach auch weiß, dass es die Effekte, Intuition usw. gibt. […] Dass ich zum Beispiel auch an die verschiedenen Biasformen denke, an die Affekte und die Intuition. Also in dem Sinne ist Kritisches Denken für mich auch so ein wachsam bleiben. Und da spielt auch immer wieder die Selbstreflexion mit rein, weil jemand, der überhaupt nicht kritisch denkt, macht sich ja vielleicht gar keine Gedanken darüber und folgt einfach seiner Meinung, die vielleicht durch Affekte geprägt ist, und dann ist das kein Kritisches Denken mehr, also weil derjenige nicht den Schritt zurück

macht. Also für mich ist das alles miteinander verknüpft (S5.3-WS23-E-M01-NB13SH, Pos. 34).

[…]

Also, dass ich mich öfter hinterfragt habe, seitdem wir die Intervention hatten, von wegen, dass ich doch ein sehr intuitiver Mensch bin, mehr als ich wahrhaben wollte, das ist mir tatsächlich erst so über die Intervention rausgekommen. […] Ja, aber irgendwie schwingt bei mir immer so ein bisschen im Hinterkopf in letzter Zeit mit: Das hast du nicht überprüft. Ist das nicht deine Bubble oder vertraust du da nicht auch einfach nur wie so ein blinder Maulwurf auf bestimmte Dinge? (S5.3-WS23-E-M01-NB13SH, Pos. 40).

Eine teilnehmende Person übertrug die Erkenntnisse aus der Vermittlungseinheit zudem auf die Schulpraxis und die Zielsetzung, als Chemielehrende ein Verfügungswissen bei Lernenden zu fördern, welches diese in die Lage versetzt, fachlich, gesellschaftlich und persönlich für sich abwägen zu können:

Weil ich finde, Fakten und Hintergründe sind noch mal ein Riesenunterschied, ob ich jetzt was habe, was emotionaler Natur ist und was mich aus irgendeinem Bedürfnis heraus oder aus irgendeinem fachlichen Kontext beeinflusst, ist ein Unterschied. Ob ich jetzt das, wie zum Beispiel in Chemie, bestimmte Reaktionen verlaufen, das kann ich mir ergründen. […] Ob ich mit den Schülern jetzt eine Silberspiegelprobe mache und ich weiß, die sollte eigentlich anschlagen, so, dann kann ich diskutieren, warum die anschlägt oder nicht auf fachlicher Ebene. Aber wenn ich jetzt mit denen, ja sag ich mal gesellschaftsrelevantes Problem mir anschaue, beispielsweise haben wir das war eine UB-Stunde dazu zu Vitamin C usw. und dann zu reflektieren: Okay, wie ist es denn? Warum ist zum Beispiel so eine Werbung mit Vitamin C teilweise irreführend für Menschen, die sich nicht damit beschäftigt haben, wie Vitamin C wirkt? Und warum ist es eigentlich teilweise oder könnte man auch den Schluss zulassen, dass es Geldmacherei manchmal ist? Das ist dann auf einer ganz anderen Ebene gewesen und das fand ich, ist halt superschön gewesen, weil die Schüler da auch kritisch hinterfragt haben: Okay, wie wirkt das? Aber auch: Wie wirkt die Werbung auf uns und warum springen wir darauf an? Warum gehen wir in den Bereich rein, dass wir das so toll finden, wenn da Sachen zusätzlich drin sind? Welche menschlichen Bedürfnisse werden da angesprochen und wie werden die teilweise auch ausgenutzt? Also da steckt ja auch eine sehr starke Analyse hinter, warum man Dinge tut und warum nicht. Und das gehört für mich zu diesem kritischen Hinterfragen auch dazu, dass ich jetzt vielleicht ein paar Schüler, das werden nicht viele sein, aber ein paar habe, die sich im nächsten Produkt, was sie kaufen, Gedanken dazu machen: Warum steht da eigentlich drauf, ist Vitamin D oder ist Vitamin C drin? So, und was bedeutet das auf chemischer Ebene? Das haben wir besprochen. Und so diese Verknüpfung zwischen fachlichem und sage ich mal weltlichen Hintergründen, persönlichen Hintergründen und auch Zielen und Motiven, finde ich, gehört einfach zu diesem kritischen Hinterfragen dazu (S5.3-WS23-E-M01-VM14OL, Pos. 40).

[…]

> Und ich meine im Lehramt, wir machen uns ja super viele Gedanken: Wie fördern wir Schüler? Wie bringen wir die dazu, selber zu denken? […] Auch im Sinne Kritisches Denken ist es ja genau das, was wir auch fördern wollen bei denen. Wir wollen ja, dass die Dinge hinterfragen. […] Unser Ziel ist ja eigentlich, dass sie klüger sind als wir. Das klingt jetzt doof, aber es ist ja genau unser Ziel, dass Sie sich Gedanken machen und informieren und dass die ein Grundverständnis von Bereichen haben, worauf die aufbauen können. Wir müssen keine Experten für alles sein, aber wir müssen den Schülern vermitteln, dass sie diese Kompetenz kriegen und dass sie sich bewusst machen, was ihre Handlungen bewirkt, dass sie sich bewusst machen: Was bedeutet es für die Menschheit? Was bedeutet das für dich, für deinen Bereich, für dein Umfeld? Für all das. Also in verschiedensten Dimensionen. Was hat das für einen Einfluss und wer spielt dir da gerade rein? Wer hat gerade Einfluss auf dich? Ist das eine Industrie, die dir was Bestimmtes sagt? Ist das irgendwo ein Hintergrund, der von dir familiär kommt? Ist das dein Umfeld, was dich in einen bestimmten Diskurs reinbringt? Und das ist was, das finde ich, das gehört dazu, dass man sich einfach Gedanken darüber macht: Was löse ich bei Menschen aus? Und das ist nicht auf Lehrer bezogen, sondern auf alle Menschen. Wenn ich was reflektiere und wenn ich mich selber reflektiere, dann reflektiere ich insofern auch: Wie mache ich das für andere? Und ja, was bewirkt das bei anderen? (S5.3-WS23-E-M01-VM14OL, Pos. 44).

Diese Aussagen zeigen zum einen die Verbindungen zwischen der fachlichen, aber auch gesellschaftlichen und individuellen Dimension des Chemieunterrichts auf, wie sie insbesondere bei der Thematisierung gesellschaftlich oder persönlich relevanter Themen mit naturwissenschaftlichem Bezug auftreten. Zum anderen drücken sie eine direkte Übertragungsleistung des im universitären Kontext Erlernten in die Schulpraxis aus. Damit subsumieren diese Aussagen die Zielsetzung des Projektes und verdeutlichen dessen potenziellen Beitrag zur Bildung und Kompetenzentwicklung Lernender.

5.5.5 Abschluss der Studie S5

Die Ergebnisse der Studie S5 weisen auf eine hohe Stabilität der eingesetzten Instrumente hin, was sich in der Reproduzierbarkeit der Resultate der Studie S3 und S4 sowie in der hohen Interkoderreliabilität der Kategorien ausdrückt. Zudem zeigte sich die in Studie S4 bereits angedeutete Kontextunabhängigkeit auf qualitativer Ebene auch hier in Studie S5, in welcher fünf verschiedene Kontexte, die Kontexte A und B im Fragebogen und die Kontexte C, D und E im Interview, zum Einsatz kamen. Beide Erhebungsformen wiesen vergleichbare Ergebnisse hinsichtlich der Veränderungen der hinzugenommenen Bewertungsperspektiven sowie des Verständnisses relevanter Begrifflichkeiten auf. Trotz der erprobten Möglichkeit, die Bewertungsstrukturen auch mittels eines Fragebogens bei einer breiteren Stichprobe erheben zu können, muss an dieser Stelle hervorgehoben werden, dass die offenen

kontextualisierten Fragestellungen des Fragebogenformats nicht an die Tiefe eines Interviews heranreichen können. Die Stärke des Instruments kommt aus diesem Grund nur in Kombination aus Fragebogen- und Interviewerhebung zum Tragen. Diese Verknüpfung ermöglicht damit einerseits den Abgleich zwischen der im Fragebogen erhobenen subjektiven Selbsteinschätzung hinsichtlich der Reflexion verschiedener Einflussfaktoren und deren Berücksichtigung im Bewertungsprozess sowie der durch das Interview vorgenommenen Fremdeinschätzung, inwieweit unterschiedliche Denkperspektiven tatsächlich in den zu evaluierenden Kontexten herangezogen werden. Andererseits trägt sie zur konvergenten Validität bei, indem die Ergebnisse der offenen Fragen des Fragebogens mit den Interviewdaten abgeglichen werden können (Hussy et al., 2010). Dadurch wird die Aussagekraft der offenen Fragestellungen für nicht interviewte Teilnehmende erhöht und eine Erweiterung der Stichprobe ermöglicht.

Die Resultate der einzelnen Instrumente zeigen eine deutliche Veränderung hinsichtlich der zur Bewertung hinzugezogenen Bewertungsperspektiven, nicht nur auf selbstreflexiver, sondern auch auf fachlicher und epistemologischer Ebene, welche mit Bezug zum Triadischen Modell die Kernelemente Kritischen Denkens bilden. Zudem war gleichzeitig eine erhöhte subjektive Tendenz, selbstreflexives Verhalten in fachliche Bewertungen einzubeziehen, erkennbar und das Bewusstsein für Einflussfaktoren stieg. Diese Ergebnisse ließen sich sowohl qualitativ erkennen als auch quantitativ bestätigen.

In Konsequenz lässt sich daraus ableiten, dass die Vermittlungseinheit, die darauf abzielte, Kritisches Denken in konkreten Anforderungssituationen zu fördern, dieses Ziel auch im Sinne der Multiperspektivität erreichte. Rekurrierend auf die entwickelte Definition Kritischen Denkens konnten sowohl erweiterte Sichtweisen auf kognitiver Ebene, repräsentiert durch die fachlichen und epistemologischen Aspekte, als auch eine verstärkte Berücksichtigung der affektiven Ebene, ausgedrückt durch die selbstreflexiven Aspekte, erreicht werden, was der angestrebten Förderung Kritischen Denkens entspricht.

Damit kann eine Passung zwischen dem Instrument und dem definierten Konstrukt Kritischen Denkens festgestellt werden, die sich in mehr als einer Studie bestätigte. Um die Stabilität des Instruments noch weiter zu prüfen, wurden darüber hinaus studien- und stichprobenübergreifende Vergleichstests durchgeführt und die Gruppen auf signifikante Unterschiede getestet. Diese Ergebnisse werden in Kapitel (5.6) vorgestellt.

5.6 Studienübergreifende Gruppenvergleiche

Aufbauend auf der Analyse der einzelnen Studien wurden zusätzlich studien- und stichprobenübergreifende Analysen durchgeführt, um die Replikationsfähigkeit und damit die methodische Robustheit des eingesetzten Instruments weiter zu überprüfen. Die Analyse erfolgte mittels des Kruskal-Wallis-Tests, welcher eine vergleichende Untersuchung über mehrere Stichproben hinweg ermöglicht (vgl. Kap. 4.3.2.3). Als nichtparametrisches Testverfahren eignet sich der Kruskal-Wallis-Test insbesondere für ordinalskalierte Daten und setzt keine Normalverteilung voraus. Zudem erlaubt er die gleichzeitige Analyse mehrerer Gruppen mit unterschiedlichen Stichprobengrößen und gewährleistet eine vergleichbare statistische Betrachtung. Diese Zusammenführung der Stichproben steigert die statistische Power und damit die Aussagekraft, da größere Stichproben präzisere Schätzungen ermöglichen, was wiederum die Wahrscheinlichkeit erhöht, tatsächlich vorhandene Effekte zu identifizieren. Die Überprüfung der Replikationsfähigkeit stärkt zudem die methodische Robustheit, was wiederum die externe Validität der Befunde absichert. Da in allen Stichproben ausschließlich Chemielehramtsstudierende untersucht und in den verglichenen Gruppen dieselben Testinstrumente eingesetzt wurden, war eine studienübergreifende Überprüfung mittels des Kruskal-Wallis-Tests methodisch gerechtfertigt.

Im Folgenden werden die Ergebnisse des Kruskal-Wallis-Tests für die Vergleiche der Pre-Tests der Fragebogenerhebungen vorgestellt (Tabelle 141):

Tabelle 141: Ergebnisse des Kruskal-Wallis-Tests für die Pre-Tests der Fragebogenerhebungen

Kruskal-Wallis-Test							
Verglichene Gruppen	Stichproben-größe	Item	Asymp. Signifikanz	Kruskal-Wallis-H	Freiheits-grade df	Effekt-stärke f	Signifikanz
	N = 67	CRT1	0,968	0,552	4	0,056	Nicht signifikant
		CRT2	0,973	0,501	4	0,057	Nicht signifikant
		CRT3	0,940	0,791	4	0,052	Nicht signifikant
	N = 69	SR3	0,057	9,178	4	0,081	Nicht signifikant
		SR5	0,662	2,402	4	0,025	Nicht signifikant
S2.F1		SR6	0,632	2,571	4	0,022	Nicht signifikant
S3.F1		SR8	0,748	1,933	4	0,032	Nicht signifikant
S4.F1		SR9	0,413	3,946	4	0,001	Nicht signifikant
S4.F2		SR10	0,350	4,438	4	0,007	Nicht signifikant
S5.F1		IN2	0,839	1,432	4	0,040	Nicht signifikant
		IN4	0,061	9,022	4	0,079	Nicht signifikant
		IN5	0,480	3,489	4	0,008	Nicht signifikant
		IN6	0,867	1,267	4	0,043	Nicht signifikant
		IN7	0,063	8,926	4	0,077	Nicht signifikant
		IN8	0,272	5,153	4	0,018	Nicht signifikant
	N = 59	IS1	0,434	2,736	3	0,005	Nicht signifikant
		IS2	0,612	1,816	3	0,022	Nicht signifikant
		IS3	0,725	1,318	3	0,031	Nicht signifikant
		IS4	0,801	1,001	3	0,036	Nicht signifikant
		IS5	0,491	2,417	3	0,011	Nicht signifikant
		EF2	0,736	1,271	3	0,031	Nicht signifikant
		EF4_1	0,583	1,947	3	0,019	Nicht signifikant
		EF4_2	0,780	1,087	3	0,035	Nicht signifikant
		EF4_3	0,361	3,206	3	0,004	Nicht signifikant
		EF4_4	0,007	12,119	3	0,168	Signifikant
		EF4_5	0,036	8,514	3	0,101	Signifikant
		EF4_6	0,516	2,280	3	0,013	Nicht signifikant
		EF4_7	0,913	0,526	3	0,045	Nicht signifikant
		EF4_8	0,881	0,667	3	0,042	Nicht signifikant
		EF4_9	0,250	4,104	3	0,020	Nicht signifikant
S3.F1		EF4_10	0,148	5,346	3	0,043	Nicht signifikant
S4.F1		EF4_11	0,032	8,830	3	0,107	Signifikant
S4.F2		EF4_12	0,262	3,994	3	0,018	Nicht signifikant
S5.F1		EF4_13	0,306	3,619	3	0,011	Nicht signifikant
		EF4_14	0,584	1,944	3	0,019	Nicht signifikant
		EF4_15	0,668	1,562	3	0,026	Nicht signifikant
		EF4_16	0,748	1,219	3	0,032	Nicht signifikant
		EF4_17	0,398	2,960	3	0,001	Nicht signifikant
		EF4_18	0,213	4,495	3	0,027	Nicht signifikant
		EF4_19	0,716	1,356	3	0,030	Nicht signifikant
		EF4_20	0,853	0,786	3	0,040	Nicht signifikant
		EF4_21	0,083	6,673	3	0,067	Nicht signifikant
		EF4_22	0,544	2,138	3	0,016	Nicht signifikant
		EF4_23	0,766	1,144	3	0,034	Nicht signifikant
		EF4_24	0,992	0,098	3	0,053	Nicht signifikant
		EF4_25	0,195	4,700	3	0,031	Nicht signifikant
		EF4_26	0,548	2,122	3	0,016	Nicht signifikant
		BG1	0,068	7,142	3	0,076	Nicht signifikant
		BG2	0,514	2,290	3	0,013	Nicht signifikant

S4.F1 S4.F2 S5.F1	N = 43	VS	0,671	0,798	2	0,030	Nicht signifikant
		VK	0,500	1,387	2	0,015	Nicht signifikant
		BF_WI	0,747	0,583	2	0,035	Nicht signifikant
		BF_EP	0,445	1,619	2	0,010	Nicht signifikant
		BF_SR	0,051	5,971	2	0,100	Nicht signifikant
		BF_SR_IF	0,090	4,822	2	0,071	Nicht signifikant
		BF_SR_EF	0,969	0,063	2	0,048	Nicht signifikant
		KV	0,613	0,979	2	0,026	Nicht signifikant

Der CRT nach Frederick (2005) und der SRIS nach Silvia (2021) wurden in fünf Studien (S2.F1, S3.F1, S4.F1, S4.F2 und S5.F1) durchgeführt. Die Gesamtstichprobe der SRIS betrug $N = 69$, während die des CRT bei $N = 67$ lag. Diese Differenz war auf die Exklusion zweier Teilnehmender der Studie S3.F1 zurückzuführen, welche in Studie S2.F1 bereits teilgenommen und die Fragen des CRT beantwortet hatten. Keines der Items des CRT oder SRIS wies signifikante Unterschiede zwischen den Gruppen auf. Die unterschiedliche Studienprogression der Teilnehmenden (S2 - S4 im Bachelorstudium, S5 im Masterstudium) machte keine messbaren Auswirkungen auf die Erstimpulsinhibition und die Selbsteinschätzung der Tendenz, selbstreflexives Verhalten zu zeigen und daraus Einsichten gewinnen zu können, sichtbar.

Die Itemblöcke zur *Introspektion* (IS), zu den *Einflussfaktoren* (EF) und zur *Bewertung und Gewichtung* (BG) wurden ab der Studie S3.F1 eingesetzt. Insgesamt wurden vier Studien miteinander verglichen (S3.F1, S4.F1, S4.F2 und S5.F1). Die Gesamtstichprobe betrug dabei $N = 59$. Die Analyse ließ in den beiden Itemblöcken *Introspektion* (IS) sowie *Bewertung und Gewichtung* (BG) keine signifikanten Unterschiede zwischen den Gruppen erkennen. Auch die Items des Blocks zur Reflexion der *Einflussfaktoren* (EF) zeigten keine messbaren Differenzen, mit Ausnahme der Items zur Reflexion der *eigenen Grundeinstellung* (EF4_4), der *eigenen Bildung* (EF4_5) und *der eigenen Kultur* (EF4_11), welche zwar das Signifikanzniveau von $p = 0,05$ erreichten, aber nur sehr geringe Effektstärken aufwiesen.

Zusammenfassend konnten für die geschlossenen Items keine oder nur sehr geringe Gruppenunterschiede zwischen den Stichproben identifiziert werden. Die Gruppen wiesen demnach eine hohe Homogenität hinsichtlich ihrer Gewichtung und Reflexion von Einflussfaktoren, ihrer Tendenz, selbstreflexives Verhalten zu zeigen, und ihrer Erstimpulsinhibition auf. Die Studienprogression schien somit keinen erkennbaren Einfluss auf diese Merkmale zu haben.

In der Studie S4 wurden die offenen Fragen erstmals um eine Kontextualisierung erweitert. Die Gesamtstichprobe umfasste $N = 43$ Teilnehmende, die sich auf die

drei Stichproben S4.F1, S4.F2 und S5.F1 verteilten. Auch die Ergebnisse der offenen Fragestellungen ließen keine signifikanten Unterschiede zwischen den Gruppen erkennen. Daraus lässt sich ableiten, dass die Gruppen hinsichtlich ihres Vorverständnisses zu den Begrifflichkeiten *Selbstreflexion* (VS), *Kritisches Denken* (VK) und *kognitive Verzerrungen* (KV) sowie den zur Bewertung hinzugezogenen *Bewertungsaspekten* (BF) vergleichbar waren und weder die Studienprogression noch die verschiedenen eingesetzten Kontexte messbare Auswirkungen hatten.

Die Ergebnisse der Gruppenvergleiche der Pre-Fragebogenerhebungen verdeutlichten eine hohe Konsistenz über die verschiedenen Gruppen hinweg, was auf eine Stabilität und Robustheit der Messungen hinweist. Dies deutet darauf hin, dass die Erfassung der Zielkonstrukte unabhängig von externen Faktoren, wie Studiensemester, Vorwissen oder verwendeten Kontexten, erfolgt. Diese Stabilität validiert die Ausgangswerte für die bereits durchgeführten intragruppalen Pre-Post-Vergleiche und stützt die Hypothese, dass die beobachteten Veränderungen innerhalb der Studien eher auf die Vermittlungseinheit als auf Messfehler oder unsystematische Schwankungen zurückzuführen sind. Darüber hinaus stärkt dies sowohl die interne Validität der Untersuchung, also die Zuverlässigkeit innerhalb der Studie, als auch die externe Validität, indem die potenzielle Übertragbarkeit auf andere Stichproben verbessert wird.

Die Ergebnisse konnten zusätzlich durch den Vergleich der beiden Post-Erhebungen der Studien S3 und S5 bestätigt werden. Da in den Studien zwei Pre-Post-Tests durchgeführt wurden, wurde der Mann-Whitney-U-Test zur Überprüfung der Unterschiede im Post-Test verwendet (Tabelle 142).

Tabelle 142: Ergebnisse des Mann-Whitney-U-Tests für die Post-Tests der Fragebögen aus S3.F2 und S5.F2

Mann-Whitney-U-Test				
	Exakte Signifikanz zweiseitig	Z-Wert	Effektstärke	Signifikanz
IS1	0,476	-0,894	0,169	Nicht signifikant
IS2	0,675	-0,394	0,074	Nicht signifikant
IS3	0,434	-1,072	0,203	Nicht signifikant
IS4	0,868	-0,33	0,062	Nicht signifikant
IS5	0,771	-0,319	0,060	Nicht signifikant
EF2	0,509	-0,679	0,128	Nicht signifikant
EF4_1	0,952	-0,129	0,024	Nicht signifikant
EF4_2	0,960	-0,149	0,028	Nicht signifikant
EF4_3	0,573	-0,53	0,100	Nicht signifikant
EF4_4	0,898	-0,109	0,021	Nicht signifikant
EF4_5	0,023	-2,51	0,474	Signifikant
EF4_6	0,787	-0,299	0,057	Nicht signifikant
EF4_7	0,813	-0,24	0,045	Nicht signifikant
EF4_8	0,506	-0,7	0,132	Nicht signifikant
EF4_9	0,600	-0,602	0,114	Nicht signifikant
EF4_10	0,052	-1,978	0,374	Nicht signifikant
EF4_11	0,110	-1,65	0,312	Nicht signifikant
EF4_12	0,252	-1,212	0,229	Nicht signifikant
EF4_13	0,021	-2,357	0,445	Signifikant
EF4_14	0,808	-0,306	0,058	Nicht signifikant
EF4_15	0,355	-1,08	0,204	Nicht signifikant
EF4_16	0,341	-0,966	0,183	Nicht signifikant
EF4_17	1,000	0	0,000	Nicht signifikant
EF4_18	0,656	-0,511	0,097	Nicht signifikant
EF4_19	0,612	-0,543	0,103	Nicht signifikant
EF4_20	0,866	-0,195	0,037	Nicht signifikant
EF4_21	1,000	0	0,000	Nicht signifikant
EF4_22	0,399	-0,908	0,172	Nicht signifikant
EF4_23	0,212	-1,271	0,240	Nicht signifikant
EF4_24	0,973	-0,075	0,014	Nicht signifikant
EF4_25	0,655	-0,495	0,094	Nicht signifikant
EF4_26	0,180	-1,405	0,266	Nicht signifikant
BG1	0,252	-1,236	0,234	Nicht signifikant
BG2	1,000	-0,051	0,010	Nicht signifikant
SR3	0,074	-1,791	0,338	Nicht signifikant
SR5	0,786	-0,341	0,064	Nicht signifikant
SR6	0,918	-0,146	0,028	Nicht signifikant
SR8	0,989	-0,076	0,014	Nicht signifikant
SR9	0,489	-0,696	0,132	Nicht signifikant
SR10	0,547	-0,642	0,121	Nicht signifikant
IN2	0,150	-1,458	0,276	Nicht signifikant
IN4	0,562	-0,596	0,113	Nicht signifikant
IN5	0,236	-1,219	0,230	Nicht signifikant
IN6	0,253	-1,164	0,220	Nicht signifikant
IN7	1,000	0	0,000	Nicht signifikant
IN8	0,357	-0,943	0,178	Nicht signifikant

Im Vergleich der Post-Tests waren ebenfalls keine signifikanten Unterschiede zu verzeichnen, mit Ausnahme der Items zur *Reflexion der eigenen Bildung* (EF4_5) und des *medialen Kontexts* (EF4_13) im Itemblock zu den reflektierten *Einflussfaktoren* (EF), die beide in der Stichprobe der Studie 5 eine größere Berücksichtigung erfuhren. Dadurch wird die Annahme weiter gestützt, dass das Instrument konsistent über die verschiedenen Studienabschnitte hinweg misst.

Auch für die Interviewerhebungen S3.I1&2 sowie S5.I1&2 wurden jeweils für die Pre- und Post-Erhebungen Vergleichstests mittels des Mann-Whitney-U-Tests durchgeführt (Tabelle 143). Hier waren im Pre-Test Unterschiede in den Ausgangsvoraussetzungen hinsichtlich der Bewertung von Quellen erkennbar, welche auf die Studienprogression zurückzuführen sein könnten.

Im Post-Test zeigte sich in der Gruppe der Masterstudierenden in Studie S5 eine stärkere Berücksichtigung *externer Faktoren* (SR-A-EF), die sich in einem Gruppenunterschied hinsichtlich der bei der Bewertung herangezogenen *selbstreflexiven Aspekte* (SR-A) widerspiegelte. Zudem wies die Mastergruppe aus Studie S5 ein umfassenderes *Verständnis der Selbstreflexion* (SR-V) im Vergleich zur anderen Gruppe auf. Da in der Pre-Erhebung keine signifikanten Unterschiede in diesen Merkmalen festgestellt wurden, kann die beobachtete Differenz im Post-Test nicht auf vorbestehende Unterschiede im Reflexionsvermögen zwischen Bachelor- und Masterstudierenden zurückgeführt werden. Dies legt nahe, dass die beobachteten Unterschiede primär auf die Vermittlungseinheit zurückzuführen sind, die bei den Masterstudierenden einen stärkeren Effekt zeigte. Mögliche Erklärungen hierfür könnten eine höhere kognitive Anschlussfähigkeit sein, die die Integration der Inhalte der Vermittlungseinheit in bestehende Wissensstrukturen erleichtert, eine erhöhte Sensibilität für die Wege der naturwissenschaftlichen Erkenntnisgewinnung oder ein besseres epistemologisches Verständnis, das durch die verstärkte Thematisierung NOS-bezogener Inhalte innerhalb des Studiums gefördert wird. Damit scheint die Studienprogression keine messbaren Auswirkungen auf die Ausgangsvoraussetzungen der Vermittlung Kritischen Denkens zu haben, sondern eher die Effektivität der Vermittlungseinheit zu beeinflussen. Fachwissen und die Fähigkeit, es sich anzueignen und einzuordnen, stellen damit keine hinreichenden Bedingungen für Kritisches Denken dar, scheinen aber notwendige Bedingungen und unterstützende Faktoren für dessen Ausprägung zu sein.

Tabelle 143: Ergebnisse des Mann-Whitney-U-Tests für die Interviewerhebungen aus S3.I1&2 und S5.I1&2

Mann-Whitney-U-Test **Pre**

Teststatistiken[a]

	NW_A_Pre	SR_A_Pre	EP_A_Pre	KD_V_Pre	SR_V_Pre	SR_A_IF_Pre	SR_A_EF_Pre
Mann-Whitney-U-Test	31,5	28,5	0	27	27,5	28,5	35
Wilcoxon-W	86,5	83,5	55	82	82,5	83,5	63
Z	-0,384	-0,652	-3,509	-0,988	-0,774	-0,651	0
Asymp. Sig. (2-seitig)	0,701	0,514	<,001	0,323	0,439	0,515	1
Exakte Sig. [2*(1-seitige Sig.)]	,740b	,536b	<,001b	,475b	,475b	,536b	1,000b
Exakte Sig. (2-seitig)	0,753	0,538	<,001	0,593	0,390	0,550	1,000
Exakte Sig. (1-seitig)	0,383	0,269	<,001	0,314	0,213	0,257	0,556
Punkt-Wahrscheinlichkeit	0,065	0,009	0	0,255	0,036	0,01	0,108

a. Gruppenvariable: Studie
b. Nicht für Bindungen korrigiert.

Effektstärke	0,093	0,158	0,851	0,240	0,188	0,158	0,000

Mann-Whitney-U-Test **Post**

Teststatistiken[a]

	NW_A_Post	SR_A_Post	EP_A_Post	KD_V_Post	SR_V_Post	SR_A_IF_Post	SR_A_EF_Post
Mann-Whitney-U-Test	31,5	14	16	23,5	14	24	9,5
Wilcoxon-W	59,5	69	71	51,5	69	79	64,5
Z	-0,352	-2,076	-1,897	-1,199	-2,258	-1,128	-2,585
Asymp. Sig. (2-seitig)	0,725	0,038	0,058	0,231	0,024	0,259	0,01
Exakte Sig. [2*(1-seitige Sig.)]	,740b	,043b	,070b	,270b	,043b	,315b	,010b
Exakte Sig. (2-seitig)	0,777	0,041	0,068	0,282	0,031	0,288	0,010
Exakte Sig. (1-seitig)	0,395	0,021	0,035	0,137	0,025	0,143	0,007
Punkt-Wahrscheinlichkeit	0,015	0,007	0,009	0,065	0,022	0,032	0,005

a. Gruppenvariable: Studie
b. Nicht für Bindungen korrigiert.

Effektstärke	0,085	0,504	0,460	0,291	0,548	0,274	0,627

Zusammenfassend erwiesen sich die Instrumente als stabil und robust zum Messen der Zielkonstrukte der Multiperspektivität in Bewertungsprozessen sowie des Verständnisses relevanter Begrifflichkeiten. Dies wurde sowohl auf Ebene der Fremdeinschätzung als auch bei der Selbsteinschätzung bezüglich der Tendenz, selbstreflexives Verhalten zu zeigen und Einflussfaktoren in meinungsbildenden Prozessen zu berücksichtigen, sichtbar. Zwischen Studierenden im Bachelorstudiengang und Studierenden im Masterstudiengang konnten zudem keine Voraussetzungsunterschiede bezüglich der Denk- und Verständnisstrukturen festgestellt werden. Dies er-

härtet die Annahme, dass sich Kritisches Denken, im Sinne einer multiperspektivischen Denkpraxis, nicht automatisch mit der Studienprogression entwickelt und daher einer expliziten Thematisierung bedarf. Vorhandenes Wissen scheint als Anschlusswissen eine Rolle zu spielen, kommt aber nicht immer zum Ausdruck. An dieser Stelle muss zwischen dem vorhandenen Fachwissen und der Fähigkeit, Fachwissen selbstständig zu erwerben und zu integrieren, differenziert werden. Die Ergebnisse der Post-Erhebungen bekräftigten den Eindruck, dass alle untersuchten Gruppen von der Vermittlungseinheit profitierten und das Instrumentarium sensitiv genug war, die Veränderungen innerhalb der Gruppe zu messen.

Abschließend kann festgehalten werden, dass das entwickelte Instrumentarium durch die Konsistenz der Daten eine hohe interne Validität zeigte und sich in der Messung der Zielkonstrukte als stabil, robust und reliabel erwies. Gleichzeitig erzielten die Instrumente über verschiedene Stichproben hinweg vergleichbare Ergebnisse, was die externe Validität bestätigte. Die signifikanten Pre-Post-Unterschiede stützten darüber hinaus die Sensitivität der eingesetzten Instrumente, während das Fehlen systematischer Verzerrungen die methodische Robustheit zusätzlich bestärkte.

5.7 Zusammenführung der Ergebnisse

Die Zielsetzung des empirischen Teils dieser Arbeit stand unter der Forschungsfrage:

Wie kann Kritisches Denken als Konstrukt im Kontext der Chemiedidaktik operationalisiert werden?

Zur Näherung an diese Operationalisierung wurden im Vorfeld drei zentrale Untersuchungsfragen formuliert:

UF1. Welche verschiedenen Perspektiven nehmen angehende Chemielehrkräfte bei der Bewertung von gesellschaftlichen Fragestellungen mit naturwissenschaftlichem Bezug (SSI) ein?

UF2. Welches Verständnis haben angehende Chemielehrkräfte von der Begrifflichkeit des Kritischen Denkens?

UF3. Inwieweit findet Selbstreflexion bei angehenden Chemielehrkräften bei der Bewertung von gesellschaftlichen Fragestellungen mit naturwissenschaftlichem Bezug statt?

Diese Untersuchungsfragen wurden ab der Studie S3 weiter ausdifferenziert, um eine umfassende Beantwortung zu ermöglichen.

In Tabelle 144 sind die untergeordneten Untersuchungsfragen zur Untersuchungsfrage 1 aufgeführt. Zur Beantwortung der Frage nach den verschiedenen Bewertungsperspektiven wurden diese sowohl in den Interview- als auch Fragebogenerhebungen erhoben und ausgewertet.

Tabelle 144: Überblick über die untergeordneten Untersuchungsfragen des Projekts und deren Zuordnung zur Zielsetzung, Untersuchungsfrage 1

UF	Form	Kürzel	Untergeordnete Untersuchungsfragen
UF1	*Interview*	NW-A	Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug *fachliche* Aspekte?
		SR-A	Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug *selbstreflexive* Aspekte?
		EP-A	Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug *epistemologische* Aspekte?
	Fragebogen	BF	Welche *Faktoren* werden von Chemielehramtsstudierenden als *beeinflussend bei der Bewertung* von Themen mit naturwissenschaftlichem Bezug wahrgenommen?

Tabelle 145 stellt die Erhebung des Verständnisses von Kritischem Denken zur Beantwortung der Untersuchungsfrage 2 in der Interview- und Fragebogenerhebung dar. Dabei wurde das Verständnis des Begriffs einmal in seiner im Bildungsdiskurs gebräuchlichen adjektivischen Form und einmal als Konstruktbegriff erhoben (vgl. Kap. 5.2.1).

Tabelle 145: Überblick über die untergeordneten Untersuchungsfragen des Projekts und deren Zuordnung zur Zielsetzung, Untersuchungsfrage 2

UF	Form	Kürzel	Untergeordnete Untersuchungsfragen
UF2	*Interview*	KD-V	Welches *Verständnis* haben Chemielehramtsstudierende von der Formulierung *kritisch denken* oder *kritisch sein*?
	Fragebogen	VK	Welches *Verständnis* haben Chemielehramtsstudierende von dem Begriff des *Kritischen Denkens?*

Da die Selbstreflexion eine ergänzende Dimension zum Fachwissen und epistemologischen Verständnis bildet (vgl. Kap. 3.4), bislang aber im Kontext der Naturwis-

senschaften nur unzureichend untersucht wurde, war sie von besonderem Erkenntnisinteresse und erforderte eine differenziertere Betrachtung (Tabelle 146). In der Folge wurde zur Beantwortung der Untersuchungsfrage 3 die Selbstreflexion sowohl in den Interviewerhebungen als auch in den Fragebogenerhebungen beleuchtet, wobei im Fragebogen verschiedene offene und geschlossene Fragestellungen verwendet wurden, um eine sowohl breit angelegte als auch tiefe Analyse zu ermöglichen. Überdies wurden über die Erhebung selbstreflexive Bewertungsaspekte zur Beantwortung von Untersuchungsfrage 1 erfasst.

Tabelle 146: Überblick über die untergeordneten Untersuchungsfragen des Projekts und deren Zuordnung zur Zielsetzung, Untersuchungsfrage 3 (Itemblöcke mit geschlossenen Fragestellungen wurden mit einem Sternchen * markiert)

UF	Form	Kürzel	Untergeordnete Untersuchungsfragen
UF3	*Interview*	SR-V	Welches *Verständnis* haben Chemielehramtsstudierende von dem Begriff der *Selbstreflexion*?
	Fragebogen	VS	Welches *Verständnis* haben Chemielehramtsstudierende von dem Begriff der *Selbstreflexion*?
		IS*	Inwieweit reflektieren Chemielehramtsstudierende bewusst *interne Faktoren* bei der Bewertung naturwissenschaftlicher Kontexte?
		EF2*	Inwieweit sehen sich Chemielehramtsstudierende in der Lage, vornehmlich auf *fachlicher Ebene* zu *urteilen*?
		EF3	Welches *Verständnis* haben Chemielehramtsstudierende von dem Begriff der *kognitiven Verzerrungen*?
		EF4*	Welche *Faktoren* werden von Chemielehramtsstudierenden bei der Bewertung von naturwissenschaftlichen Themen besonders *reflektiert*?
		BG*	Inwieweit *berücksichtigen* Chemielehramtsstudierende *interne und externe Einflussfaktoren* bei der Bewertung naturwissenschaftlicher Kontexte, zusätzlich zum Fachwissen?
		SRIS*	Inwieweit zeigen Chemielehramtsstudierende die *Tendenz, selbstreflexives Verhalten* zu zeigen?
		CRT*	Inwieweit sind Chemielehramtsstudierende in der Lage, ihren *intuitiven Erstimpuls* bei Fragestellungen zu *relativieren* und eine rationale Problembetrachtung vorzunehmen?

Diese Ausdifferenzierung der Untersuchungsfragen erlaubte eine detaillierte Analyse der Teilbereiche zur Beantwortung der übergeordneten Untersuchungsfragen und trug damit zu einer umfassenderen Operationalisierung Kritischen Denkens bei.

Nachfolgend sollen die übergeordneten Forschungsfragen mit Bezug auf die Ergebnisse der durchgeführten Studien zusammenfassend beantwortet werden.

Bei der *Bewertung gesellschaftlicher Fragestellungen mit naturwissenschaftlichem Bezug* (UF1) zeigte sich, dass die Studierenden in den Pre-Untersuchungen nur wenige fachliche oder wissensbezogene Aspekte zur Positionierung hinzuzogen. Die Schwerpunkte lagen hierbei lediglich auf der Abwägung von Kosten oder der Nachhaltigkeit. Informationen mit stärker naturwissenschaftlichem oder fachlich-chemischem Bezug, wie das Verständnis der Funktionsweise oder die verwendeten Rohstoffe, wurden kaum erfragt oder erwogen. Unter der Berücksichtigung, dass es sich bei den Befragten um Chemielehramtsstudierende handelte, die Kontexte explizit naturwissenschaftlicher Natur waren und die Interviews in einem chemiedidaktischen Kontext durchgeführt wurden, verwunderte diese geringe Berücksichtigung fachlicher Inhalte. Daraus ließ sich ableiten, dass der Positionierungsprozess von anderen Faktoren gestaltet wurde, welche eher auf der affektiven Ebene zu verorten waren. Zwei Teilnehmende benannten dies auch explizit:

> Aber um jetzt eine Positionierung stattzufinden, müsste ich, würden mich die technischen Aspekte gar nicht so belangen, wie das jetzt genau funktioniert, sondern eher, da bin ich glaube ich eher ergebnisorientiert (S5.I1-WS23-D-M01-AB30VT, Pos. 40).

> Also der erste Impuls muss immer dann raus und dann geht, ist es meistens sehr emotional (S5.I1-WS23-D-M01-VF42IB, Pos. 40).

Auch epistemologische Aspekte zur Bewertung von Quellen wurden in den Pre-Erhebungen weniger berücksichtigt. Innerhalb dieser Aspekte standen sowohl bei Bachelor- als auch bei Masterstudierenden insbesondere die Herkunft der Informationen und die Übereinstimmung mit anderen Quellen im Fokus, wobei diese Aspekte bei den Masterstudierenden stärker ausgeprägt waren.

Selbstreflexive Aspekte fanden ebenfalls wenig differenziert Einzug in den Bewertungsprozess, wobei hier in den Pre-Erhebungen verstärkt selbstzentrierte Aspekte, wie die persönliche Relevanz der Thematik in der eigenen Lebensrealität, angeführt wurden. Auch der Abgleich mit dem sozialen Umfeld stellte einen der größten Einflussfaktoren zur Positionierung dar, was eine Übertragbarkeit der Annahmen der Social Norms Theory (Elster, 1989, 2008) und der Theory of Planned Behaviour (Ajzen, 1991), welche die soziale Norm als Verhaltensprädiktor herausstellen, auf die untersuchten Stichproben andeutet.

Übergreifend betrachtet wurden die Positionierungen in den Pre-Erhebungen vorwiegend wenig begründet und einseitig vorgenommen, was sich in den wenigen zur Bewertung herangezogenen Perspektiven äußerte. Gleichzeitig wurde deutlich, dass den Teilnehmenden die Positionierungen innerhalb der Kontexte aufgrund der Uneindeutigkeiten der ausgewählten SSI-Kontexte schwerfiel, was in anderen Studien ebenfalls beobachtet werden konnte (vgl. Kap. 2.1.6). Bezüglich der Studienprogression konnten im Bewertungsverhalten der Pre-Erhebungen nur geringe bis keine Unterschiede in den Ausgangsvoraussetzungen festgestellt werden, was darauf schließen lässt, dass die Aneignung von Fachwissen im Studium nicht selbstständig zur Verbesserung der Bewertungsfähigkeit und des Kritischen Denkens im Sinne einer multiperspektivischen Denkpraxis führt. Eine gezielte und explizite Förderung der Fähigkeit zum Kritischen Denken in gesellschaftlichen Kontexten mit naturwissenschaftlichem Bezug ist daher nicht nur als wünschenswert, sondern als nötig zu betrachten. Diese Förderung sollte jedoch nicht additiv, sondern integrativ zum Fachwissenserwerb gestaltet werden, da Kritisches Denken ein grundlegendes Fachwissen zur Erweiterung und Anbindung benötigt. Ein integrativer Vermittlungsansatz würde zudem die zeitlichen Ressourcen für die Vermittlung stärker fachlicher Inhalte wenig beeinflussen und sich daher für eine unterrichtspraktische Umsetzung eignen. Damit scheint die Vermittlung Kritischen Denkens vergleichbar mit der Vermittlung von NOS, für die ebenfalls ein dynamischer und integrativer Ansatz als zielführend betrachtet wird (Clough, 2006).

Das *Verständnis des Begriffs des Kritischen Denkens* (UF2) erwies sich in den Pre-Erhebungen über alle Studien hinweg als stark subjektiv geladen und lag im Wesentlichen naiv definiert vor. Am häufigsten wurde Kritisches Denken als das unkonkrete Hinterfragen von Dingen oder als eine grundsätzlich skeptische Haltung beschrieben, nicht alles zu glauben. Skepsis allein bietet jedoch nur wenig alltäglichen Nutzen in Entscheidungssituationen, insbesondere dann, wenn sie nicht präzisiert und zielgerichtet vorliegt (Collins, 2009; Sharon & Baram-Tsabari, 2020). In diesem Fall führt Skepsis eher zu einer Entscheidungspassivität, da durch die ungläubige Grundhaltung jedwede Entscheidungsgrundlage beeinträchtigt wird. Auch wenn ein Begriffsverständnis nicht zwingend notwendig ist, um eine Kompetenz zum Ausdruck bringen zu können, kann davon ausgegangen werden, dass ein adäquates Begriffsverständnis zur Verstärkung der Kompetenz beiträgt. Dies unterstreicht ebenfalls die Notwendigkeit der expliziten Thematisierung Kritischen Denkens und dessen Implementation in die Ausbildung angehender Chemielehrkräfte.

Die Fähigkeit zur Selbstreflexion charakterisiert, neben dem naturwissenschaftlichen Fachwissen und dem epistemologischen Verständnis, welches beispielsweise über die Vermittlung von NOS adressiert wird, die ergänzende Komponente im Triadischen Modell Kritischen Denkens in der Naturwissenschaftsdidaktik. Angesichts des weitreichenden Einflusses affektiver Faktoren auf Denk- und Entscheidungsprozesse ist es neben der fachlichen Vermittlung von Bedeutung, ein Bewusstsein für das eigene Denken zu schaffen und die Fähigkeit zur Metakognition zu fördern (Gigerenzer, 1991; Greenhalgh, 2002; Maddux, 1995; Taylor et al., 2009). Diese Fähigkeit zur *Selbstreflexion* (UF3) fand in den Pre-Erhebungen bei der Bewertung gesellschaftlich relevanter Kontexte mit naturwissenschaftlichem Bezug nur recht einseitig Anwendung. Die reflektierten Bewertungsperspektiven waren größtenteils selbstzentriert auf die Relevanz für die eigene Person ausgerichtet und die Befragten gaben an, interne Faktoren nur mäßig zu reflektieren. Auch die Einflüsse interner und externer Faktoren auf Bewertungsprozesse wurden nur mäßig bis schwach eingeschätzt und infolgedessen in diesen wenig berücksichtigt. Die Pre-Erhebungen ließen zudem über alle Studien hinweg erkennen, dass die Teilnehmenden eine geringe bis mäßige Tendenz hatten, selbstreflexives Verhalten zu zeigen, und ebenso nur eine geringe bis mäßige Erstimpulsinhibition aufwiesen. Das Verständnis der Selbstreflexion lag indes überwiegend auf einem informierten Niveau vor, wobei der Begriff von der Mehrheit als das Überdenken von Handlungen definiert wurde. Introspektive Prozesse oder das Bewusstsein für externe Prägungen wurden hingegen weniger zur Selbstreflexion gezählt.

Mit Blick auf die hohe Relevanz dieser metakognitiven Fähigkeit in Bewertung- und Entscheidungsprozessen zeugt dessen geringer Einbezug seitens der Teilnehmenden von Defiziten, die einer Adressierung bedürfen. Wenn fachliche Aspekte wenig zur Positionierung angeführt oder erfragt und potenzielle Einflussfaktoren nur unzureichend reflektiert werden, liegen die getroffenen Entscheidungen potenziell in der affektiven Ebene begründet. Ohne eine aktive Reflexion verbleiben die internen und externen Einflüsse im Unbewussten und leiten die Entscheidungen intuitiv an. Infolgedessen steigt die Wahrscheinlichkeit kognitiver Verzerrungen und möglicher Fehlurteile, was insbesondere in komplexen und lebensweltlich relevanten Kontexten auf individueller und gesellschaftlicher Ebene negative Auswirkungen zur Folge haben kann.

Sowohl die qualitativen und quantitativen Befunde der Pre-Post-Studien als auch die Studierendenrückmeldungen zur Vermittlungseinheit unterstützen jedoch die Annahme, dass Kritisches Denken, welches neben fachlichen auch epistemologische

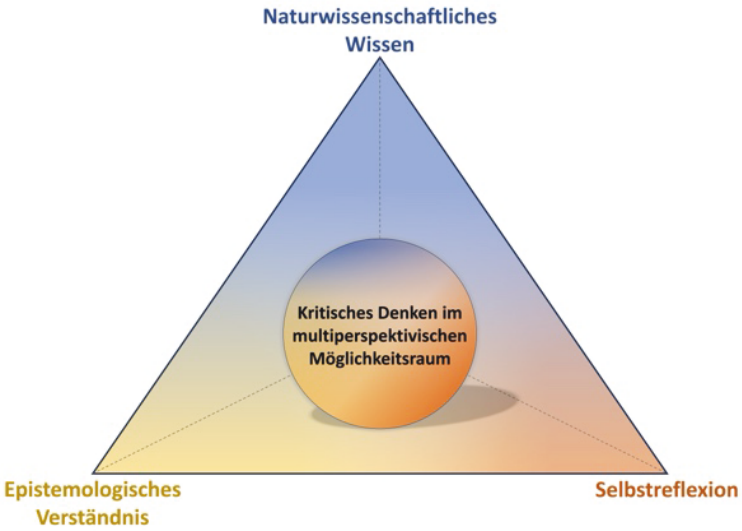
und selbstreflexive Perspektiven beinhaltet, grundsätzlich adressierbar und förderbar ist. Die Post-Erhebungen nach der Durchführung der Vermittlungseinheit ließen fundiertere Verständnisse und signifikante Perspektiverweiterungen in konkreten Anforderungssituationen erkennen, welche auch nachhaltig Bestand hatten. Dabei war gleichzeitig eine gesteigerte subjektive Tendenz zu selbstreflexivem Verhalten und ein erhöhtes Bewusstsein für beeinflussende interne und externe Faktoren zu verzeichnen, was auf ein höheres Selbstbewusstsein im Umgang mit Einflussfaktoren hindeutet. Sowohl die Ergebnisse der Selbsteinschätzung als auch der Fremdeinschätzung zeigten signifikante Verbesserungen im reflexiven und multiperspektivischen Denken, was die Annahme einer grundsätzlichen Veränderbarkeit beider Dimensionen stützt.

Die erprobte Methode der Systemisch-reflexiven Stellungnahme, welche konkrete Anforderungssituationen zur Auseinandersetzung mit einem selbst ausgewählten gesellschaftlichen Kontext mit naturwissenschaftlichem Bezug schaffte, wurde seitens der Studierenden sehr positiv aufgenommen, was sich auch in der Mitarbeitsbereitschaft und den erhaltenen Rückmeldungen widerspiegelte.

Aus methodischer Perspektive erwies sich der Einsatz von Interviews zur Fremdeinschätzung der Bewertungsstrategien als gewinnbringend und zielführend, da viele bestehende Instrumente versuchen, komplexe Konstrukte, wie die Selbstreflexion, ausschließlich über Selbsteinschätzungen zu erschließen. Diese sind jedoch anfällig für systematische Verzerrungen, weshalb eine Ergänzung durch die Fremdperspektive methodisch notwendig erscheint.

Die größte Herausforderung im Rahmen der Fremdeinschätzung bestand im Aufbau der Erhebungsmethode, da das Instrument eine Unabhängigkeit vom tatsächlich vorhandenen Fachwissen gewährleisten musste. Andernfalls hätte das Fachwissen den Fokus der Erhebung dominiert, anstatt die Multiperspektivität Kritischen Denkens zu erfassen. Durch die Ermöglichung der Fragenformulierung für die Teilnehmenden und die Fokussierung auf den Prozess der Entscheidungsfindung selbst, anstelle des Ergebnisses, konnten die eigentlichen Denkperspektiven untersucht werden, was eine Operationalisierung Kritischen Denkens ermöglichte. Gleichzeitig machten diese Unabhängigkeit vom Fachwissen und die Prozessorientierung die Verwendung verschiedener Kontexte möglich und ließen so die Untersuchung Kritischen Denkens in verschiedenen Bewertungsszenarien zu. In Anlehnung an die konzeptionellen Ansätze und die Entwicklung des *Dual-Process Models of Critical Thinking* (DP-CT) sowie des *Triadischen Modells Kritischen Denkens* erlaubte diese Methodik

nicht nur die systematische Erhebung der Denkperspektiven, sondern auch deren direkte konzeptionelle Einordnung in das Triadische Modell selbst (Abbildung 31).



S5.I1&2 - Übersicht Follow Up, % max. Kat.

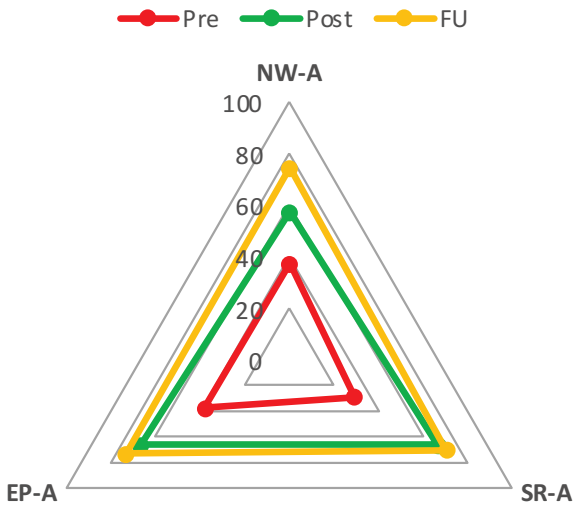


Abbildung 31: Übertragbarkeit der Ergebnisse auf das Triadische Modell Kritischen Denkens

Das semistrukturierte Leitfadeninterviewformat ermöglichte überdies eine Fokussierung auf die intendierten Schwerpunkte und zugleich eine flexible Gestaltung des Interviewverlaufs, was Nachfragen erlaubte und damit eine tiefergehende Erhebung förderte.

Die Hinzunahme des Fragebogenformats gestattete darüber hinaus den Abgleich der Fremdeinschätzung mit der Selbsteinschätzung der Studierenden und dadurch weiterführende Einblicke in deren selbstreflexives Verhalten. Ergänzend dazu verstärkte die Integration bereits etablierter Werkzeuge, wie des CRT nach Frederick (2005) oder der SRIS nach Silvia (2021), triangulierend die Aussagekraft der selbsterstellten Itemblöcke und deren Verlässlichkeit. Das Erweitern der offenen Fragestellungen

des Fragebogens um gezielte Kontexte ließ zusätzlich zum Interview die Erhebung von Bewertungsperspektiven bei einer größeren Stichprobe zu. Diese erreichte wegen des limitierten Erhebungsformats nicht die Tiefe der Interviewerhebungen, ließ aber zentrale Tendenzen erkennen, welche durch den Abgleich mit den Ergebnissen der geführten Interviews an Aussagekraft gewannen.

Durch die sukzessive Ausschärfung der Kategorien über die ersten drei Studien hinweg erreichten diese eine hohe Trennschärfe, die sich ebenfalls in einer hohen Interkoderreliabilität äußerte. Diese Werte bestätigten sich nicht nur in einer Studie, sondern konsistent über zwei Interview- und drei Fragebogenerhebungen hinweg.

Obwohl verschiedene Settings für die Interviewführung und die offenen Fragen des Fragebogens verwendet wurden, konnten gruppenübergreifende Vergleiche keine messbaren Unterschiede innerhalb und zwischen den Stichproben zeigen. Diese erkennbare Kontextunabhängigkeit begünstigt damit den Einsatz in verschiedenen Anforderungssituationen, in denen gesellschaftlich relevante Kontexte mit naturwissenschaftlichem Bezug eine Rolle spielen.

In Summe betrachtet erwies sich die Kombination aus Selbst- und Fremdeinschätzung als gewinnbringende Verknüpfung unterschiedlicher Erhebungsmethoden zur Schaffung eines integrativen Instrumentariums zur Operationalisierung Kritischen Denkens. Die in der Einleitung zu Kapitel 4 beschriebenen Gelingensbedingungen wurden dabei vollumfänglich berücksichtigt und die angestrebte Zielsetzung nachhaltig realisiert.

5.8 Abschließende Diskussion der durchgeführten Studien

Die in den vorangegangenen Kapiteln vorgestellten Ergebnisse der einzelnen Studien lassen darauf schließen, dass eine Operationalisierung Kritischen Denkens über die Untersuchung der in Bewertungsprozessen hinzugezogenen Denkperspektiven sowie die Förderung Kritischen Denkens durch explizite Thematisierung umsetzbar ist. Da die gezogenen Stichproben jedoch vollständig aus Chemielehramtsstudierenden bestanden, muss auf eine gewisse Homogenität hinsichtlich des Bildungsstands und damit verbunden des Abstraktionsvermögens hingewiesen werden.

Aufgrund der kleinen Stichprobengrößen lassen sich die Ergebnisse zudem nicht ohne weiteres verallgemeinern. Da hier aber der Schwerpunkt auf der qualitativen Datenaufnahme, -auswertung und Interpretation sowie auf der Ausarbeitung von

Tendenzen lag und quantitative Methoden nur ergänzend zur Sicherstellung der Aussagekraft eingesetzt wurden, stellten die angewandten statistischen Verfahren eine methodisch angemessene Möglichkeit dar, trotz kleiner Stichproben belastbare Ergebnisse zu gewinnen.

Die Auswertung der hier gewonnenen qualitativen Daten weist zudem einen subjektiven Anteil durch die Interpretation der Kodierregeln auf, wurde jedoch durch eine möglichst regelgeleitete Vorgehensweise methodisch abgesichert. Zur Gewährleistung der methodischen Qualität kamen hierfür unter anderem die qualitative Inhaltsanalyse nach Kuckartz (2016), Interview- und Kodierleitfäden, Kodierkonferenzen in Anlehnung an Mayring (2010) und die Bestimmung der Interkoderreliabilität zum Einsatz.

Ferner lassen sich verzerrende Effekte, wie die soziale Erwünschtheit in Befragungssituationen, wie Interviews oder Fragebögen, nie vollständig ausschließen, zumal der Untersuchende in den Studien S1 und S3 auch als bewertender Dozent auftrat. Um diese Effekte gering zu halten, wurden die Pseudonymisierung und Freiwilligkeit der Teilnahme sichergestellt, die Fragen in beiden Erhebungsformen möglichst neutral formuliert und Erhebungen zusätzlich in anderen, nicht selbst gestalteten Seminaren durchgeführt. Effekte, die unmittelbar auf die Vermittlungseinheit zurückzuführen waren, wie das erweiterte Verständnis der Begriffsdefinitionen, konnten erwartungsgemäß beobachtet werden, wurden jedoch durch die Durchführung ergänzender Follow-Up-Interviews zur Prüfung der Nachhaltigkeit validiert.

Trotz der oben genannten Einschränkungen zeigen die Ergebnisse der durchgeführten Studien, dass eine Einflussnahme auf derlei komplexe Denkprozesse, wie im Kontext Kritischen Denkens, durch eine explizite Thematisierung grundsätzlich möglich zu sein scheint. Dies schafft konkrete Perspektiven für die weitere Entwicklung von Fördermöglichkeiten mit der Absicht einer langfristigen Etablierung Kritischen Denkens in der Chemielehrer*innenbildung und der Übertragung in die Schule.

6 Fazit und Ausblick

Das übergeordnete Ziel dieser Dissertation bestand darin, einen Beitrag zur langfristigen Förderung Kritischen Denkens im naturwissenschaftlichen Unterricht zu leisten. Zur Annäherung an die Beantwortung dieser Leitfrage wurde die Beantwortung zweier Forschungsfragen angestrebt. Während die konzeptionelle Forschungsfrage (1) das Wesen Kritischen Denkens zu definieren sucht, nimmt sich die empirische Forschungsfrage (2) der Operationalisierung dieses komplexen Konstruktes an, um eine Grundlage zur Förderung Kritischen Denkens zu legen.

(1) Wie kann Kritisches Denken als Konstrukt definiert werden?

(2) Wie kann Kritisches Denken im Kontext der Chemiedidaktik operationalisiert werden?

Diese Forschungsfragen werden im Folgenden abschließend beantwortet und die Implikationen der gewonnenen Erkenntnisse für die Chemiedidaktik und den naturwissenschaftlichen Unterricht dargelegt und diskutiert. Auf Grundlage dessen erfolgt schließlich ein Ausblick auf weiterführende Forschungsansätze und noch offene Forschungsdesiderate.

6.1 Beantwortung der konzeptionellen Forschungsfrage

Das Grundziel der Bildung im Allgemeinen und naturwissenschaftlicher Grundbildung im Besonderen liegt nicht allein in der inhaltlichen Vermittlung von Wissen, sondern zugleich in der Befähigung, dieses auch in lebensweltlichen Kontexten zur Orientierung nutzen zu können. Dennoch zeigt sich insbesondere in komplexen gesellschaftlichen Problemstellungen mit naturwissenschaftlichem Bezug (SSI), dass das Vorhandensein von Wissen nicht obligat auch zu adäquatem Handeln führt und stark von affektiven Faktoren beeinflusst wird. In den Bestrebungen, diese Wissens-Handlungs-Lücke zu überbrücken, kommt dem Kritischen Denken eine zentrale Bedeutung zu. Dieses wird zwar als Schlüsselkompetenz für lebenslanges Lernen benannt und auch in der Naturwissenschaftsdidaktik als relevante Fähigkeit und Lernziel formuliert, verbleibt aber dennoch ambig und uneindeutig definiert, was eine gezielte Operationalisierung und Förderung erschwert. Aus dieser Notwendigkeit heraus wurde eine systemische Neustrukturierung bestehender Begriffsdefinitionen unter Hinzunahme interdisziplinärer kognitionspsychologischer Erkenntnisse vorgenommen und eine Neudefinition Kritischen Denkens erarbeite

Kritisches Denken (KD) bezeichnet eine sorgfältige, systematische und vor allem begründete Entscheidungsfindung bei der Auseinandersetzung der*s Denkenden mit einem Objekt oder Subjekt. Diese Auseinandersetzung beinhaltet die Evaluation und Anwendung des eigenen Wissens und der eigenen Fähigkeiten auf kognitiver Ebene sowie die Berücksichtigung der eigenen Emotionen, Werte, Einstellungen und Motivationen auf affektiver Ebene. Es umfasst zudem eine konstante Selbstreflexion und -regulation, die eine Rückkopplung zur*m Denkenden selbst ermöglicht und diese*n nicht nur in die Position der*s Bewertenden, sondern auch in die Position der*s zu Bewertenden versetzt (eigene Definition u. a. in Anlehnung an Haidt, 2008; Halpern, 2014; Kahneman, 2012; D. Kuhn, 1999; Rafolt et al., 2019).

Die starke Berücksichtigung affektiver Faktoren und der besondere Stellenwert der Selbstreflexion gehen dabei insbesondere auf die Dual-Process Theory zurück, welche sich als tragfähiger theoretischer Rahmen zur Erklärung der Interdependenzen zwischen affektiv-intuitiven und rational-analytischen Denkprozessen in Bewertungskontexten erweist. Damit geht Kritisches Denken über herkömmliche Kompetenzmodelle hinaus und stellt nicht nur eine Fähigkeit, sondern den Prozess geleiteten Denkens selbst dar.

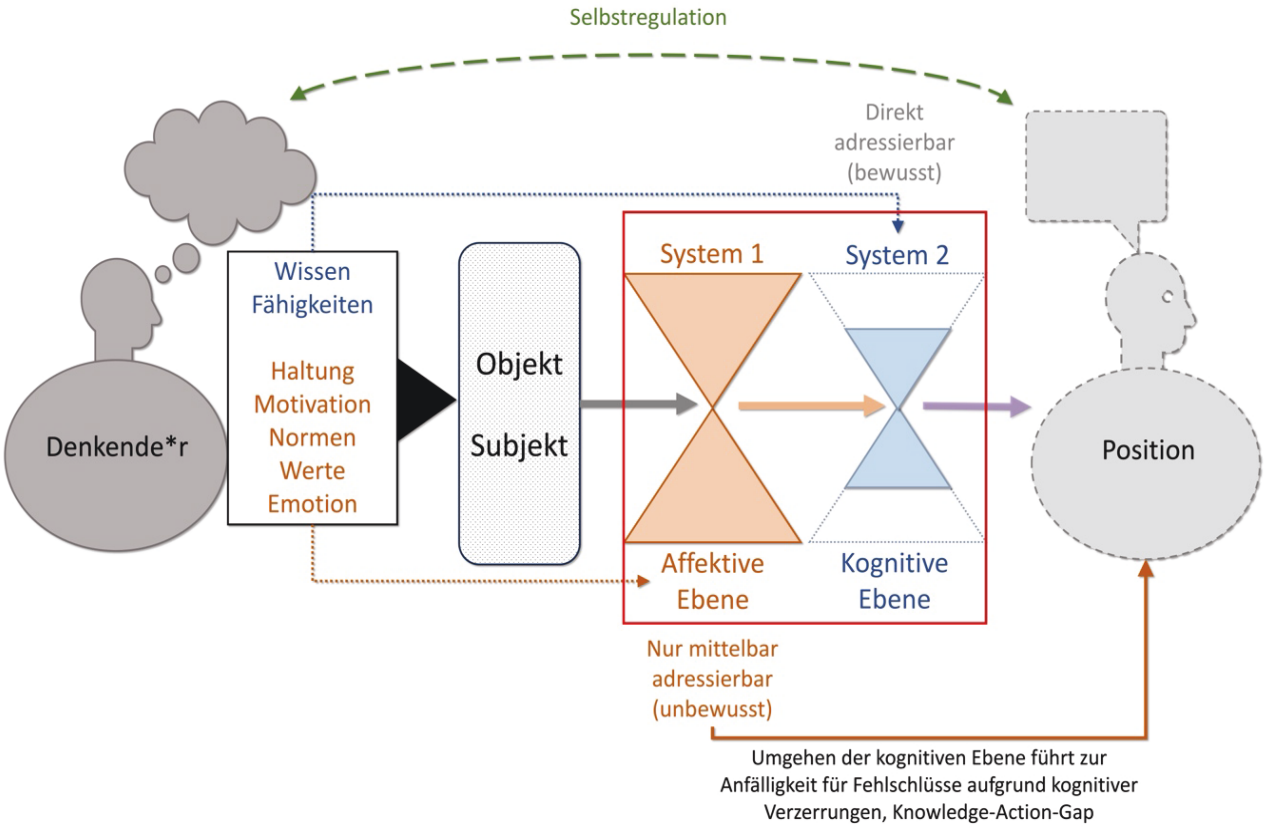


Abbildung 32: Dual-Process Model of Critical Thinking (DP-CT) (eigenes Modell)

Das *Dual-Process Model of Critical Thinking* (DP-CT) verdeutlicht den in der Definition beschriebenen Denkprozess (Abbildung 32) und die Bedeutung der Selbstreflexion als vermittelnde Komponente zwischen der affektiven und kognitiven Ebene. Da die affektive Ebene evolutiv der kognitiven Ebene vorgelagert ist, nehmen interne und externe Faktoren, wie Normen, Werte oder Emotionen, unbewusst Einfluss auf die Entscheidungsprozesse und können infolgedessen zum Umgehen der kognitiven Ebene führen. Dies begünstigt systematische Denkfehler und führt zu einer Anfälligkeit für fehlerhafte Schlussfolgerungen und Urteile. Die Selbstreflexion erlaubt die aktive Kenntnisnahme von internen und externen Einflussfaktoren und hebt diese damit auf die bewusste Ebene. Dadurch wird ein Möglichkeitsraum geschaffen, in dem diese Faktoren zum vorhandenen Wissen in ein Verhältnis gesetzt und folglich in den Bewertungsprozess miteinbezogen werden können. Wissen bleibt in diesem Kontext weiterhin als Bewertungsgrundlage von Bedeutung, muss aber vor allem bei relevanten gesellschaftlichen und individuellen Entscheidungen mit naturwissenschaftlichem Bezug durch die selbstreflexive Komponente ergänzt werden, um zum Tragen kommen zu können.

Mit Blick auf die Chemiedidaktik geht Kritisches Denken damit deutlich über die bestehende Bewertungskompetenz hinaus, welche stärker die Anwendung fachlicher Perspektiven adressiert. Diese fachlichen Blickwinkel sind indes verstärkt auf der kognitiven Ebene zu verorten und unterliegen damit den zuvor beschriebenen Einflüssen der affektiven Ebene, was die Sinnhaftigkeit einer Erweiterung unterstreicht.

Auch im Kontext der Vermittlung von naturwissenschaftlichen Arbeits- und Denkweisen in Zusammenhang mit Nature of Science (NOS) kommt Kritischem Denken eine relevante Bedeutung zu, da NOS die soziokulturelle Eingebundenheit naturwissenschaftlicher Erkenntnisse thematisiert und den Menschen als subjektives Element im Prozess der Wissensgenese wahrnimmt. Kritisches Denken erweitert diese Perspektive um die individuelle Entscheidungsebene und berücksichtigt überdies den Menschen auch als Rezipierenden naturwissenschaftlichen Wissens. Damit verknüpft kommt der Fähigkeit zur selbstständigen Wissensaneignung und -integration vor allem im Kontext der fortschreitenden Digitalisierung eine vorrangige Bedeutung zu, da Informationen über das Internet jederzeit verfügbar sind und von den Rezipierenden bewertet und reflektiert werden müssen. Kritisches Denken ermöglicht hierbei in seiner Prozessorientierung eine Abwägung zwischen dem inhaltlichen Fachwissen, dem epistemologischen Verständnis und der Selbstreflexion, um den zunehmenden Herausforderungen des digitalen Wandels zu begegnen. Das *Triadische Modell Kritischen Denkens* (Abbildung 33) verdeutlicht, dass Kritisches Denken nicht

als einzelne Fähigkeit, sondern als systemisches Konstrukt verstanden werden sollte, in welchem sich die einzelnen Bestandteile gegenseitig bedingen. Während das fachliche Wissen die notwendige inhaltliche Grundlage für den Bewertungsprozess und damit das Anschlusswissen bildet, befähigt das epistemologische Verständnis zur Aneignung, Bewertung und Integration neuer Informationen. Die Selbstreflexion schafft ergänzend ein Bewusstsein für interne und externe Einflussfaktoren und erlaubt damit die Einnahme einer objektiveren Position.

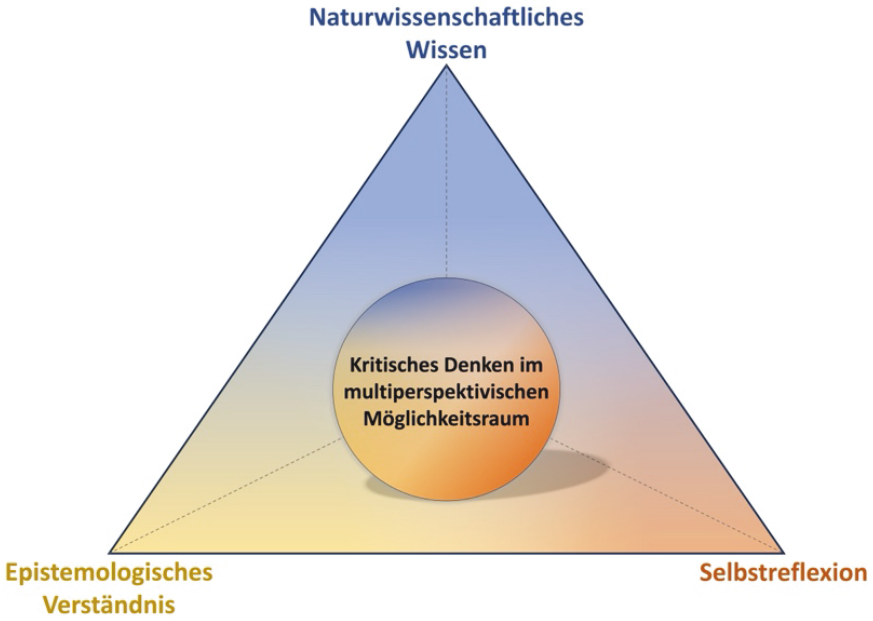


Abbildung 33: Triadisches Modell Kritischen Denkens in der Naturwissenschaftsdidaktik (eigenes Modell)

Kritisches Denken lässt sich in Folge als systemische und multiperspektivische Denkpraxis charakterisieren, die des Zusammenspiels aller drei Komponenten bedarf.

Die hier entwickelte Definition Kritischen Denkens führt damit verschiedene bestehende Definitionsansätze zusammen, indem sie sowohl kognitive als auch affektive Dimensionen integriert und metakognitive Komponenten einbezieht. Auf diese Weise wird der inhärenten Komplexität des Konstrukts angemessen Rechnung getragen, während gleichzeitig eine konzeptuelle Reduktion vorgenommen wird, die eine empirisch fundierte Operationalisierung ermöglicht, welche im Folgenden weiter ausgeführt wird.

6.2 Beantwortung der empirischen Forschungsfrage

Um Kritisches Denken langfristig gezielt fördern zu können, ist im Vorfeld eine tragfähige Operationalisierung des Konstrukts erforderlich. Diese dient nicht nur dazu, Kritisches Denken messbar zu machen und in beobachtbare Variablen zu überführen, sondern auch der Präzisierung und Schärfung seiner Definition.

Die im Rahmen der empirischen Studien durchgeführte Operationalisierung erfolgte iterativ und auf Grundlage der konzeptionellen Definition. In den Vorstudien S1 und S2 wurden hierfür verschiedene Erhebungsansätze mit Studierenden des Chemielehramts erprobt und auch bestehende psychologische Testinstrumente zur Triangulation integriert. Rekurrierend auf das Verständnis Kritischen Denkens als multiperspektivische Denkpraxis erwiesen sich insbesondere prozessorientierte Erhebungsansätze als besonders gewinnbringend. Diese erfassten die Denk- und Bewertungsperspektiven der Teilnehmenden in konkreten gesellschaftlich relevanten Bewertungskontexten mit naturwissenschaftlichem Bezug (SSI), ohne sich allein auf ihre finale Positionierung zu fokussieren. Ein zentraler Vorteil dieser Ansätze bestand darin, dass die Bewertungsperspektiven unabhängig vom spezifischen Fachwissen der Teilnehmenden erhoben werden konnten. Auf diese Weise standen das Kritische Denken und die damit verbundenen Denkperspektiven anstelle des spezifischen Fachwissens der Teilnehmenden im Vordergrund. Ein bedeutender Faktor war diesbezüglich die im Rahmen der initialen Fragengenese des Fragebogens der Studie S1 erprobte Herangehensweise, den Teilnehmenden die eigenständige Formulierung als wichtig erachteter Fragen zu ermöglichen. Diese Fokussierung legte die für die Bewertung wesentlichen Faktoren offen, ohne auf konkretes inhaltliches Fachwissen angewiesen zu sein, und differenzierte damit zwischen fachlichem Wissen und der Fähigkeit, dieses in Bewertungskontexten eigenständig zu erschließen.

Obwohl sich der grundsätzliche Erhebungsansatz der Fragengenese bereits als vielversprechend herausstellte, zeigte sich das Erhebungsmedium des Fragebogens als nicht tiefgreifend genug, weshalb dieser Ansatz in Vorstudie S2 in das Interviewformat übertragen wurde. Die Verknüpfung des prozessorientierten Erhebungsansatzes und der Interviewmethode ermöglichte Einblicke in die individuellen Herangehensweisen an Bewertungsprozesse und ließ erkennen, welche Faktoren als zentral für die Bewertung erachtet wurden. Zudem spielte die Einbettung konkreter Bewertungsanlässe eine zentrale Rolle bei der Erfassung der Bewertungsperspektiven, wobei sich weniger der fachliche Inhalt als vielmehr die Existenz eines Bewertungsanlasses als entscheidend erwies.

Ab Studie S2 wurde zusätzlich zur Fremdeinschätzung auch die Selbsteinschätzung der Studierenden mittels eines Fragebogens erhoben. Dieser fokussierte sich auf die Erhebung der Selbstreflexion, da diese die erweiternde Dimension zum Fachwissen sowie dem epistemologischen Verständnis in der Neudefinition Kritischen Denkens darstellte und daher von besonderem Erkenntnisinteresse war. Die Selbsteinschätzung eröffnete zusätzlich zur Fremdeinschätzung Einblicke in das Selbstempfinden der Teilnehmenden und differenzierte die Analyse der Selbstreflexion weiter aus. Zur Triangulation der ab Studie S3 implementierten selbsterstellten Items zur Selbstreflexion, fand die Self-Reflection and Insight Scale (SRIS) nach Silvia (2021) als etabliertes Testinstrument Einzug in den finalen Fragebogen, ebenso wie der Cognitive Reflection Test (CRT) nach Frederick (2005).

Die Ausschärfung der Definition durch die Vorstudien S1 und S2 sowie die Konkretisierung des Instruments zur Erhebung der Denkperspektiven führten in Verbindung mit den konzeptionellen Fortschritten ab Studie S3 zur Konsolidierung der Definition Kritischen Denkens. In der Folge diente die Operationalisierung zusätzlich der Überprüfung der Funktionalität und Tragfähigkeit der entwickelten Definition. Das *Triadische Modell Kritischen Denkens* bildete im Rahmen der Kodierleitfäden die Grundlage für die weitere Untersuchung der Bewertungsperspektiven, indem die übergeordneten Kategorien deduktiv aus dem Modell entlehnt und induktiv weiter ausdifferenziert wurden. Der vornehmliche Fokus der Analyse lag dabei nicht auf der Häufigkeit der Nennungen einer Kategorie, sondern darauf, wie viele differente Perspektiven bei der Bewertung hinzugezogen und eingenommen wurden, was dem multiperspektivischen Charakter Kritischen Denkens entspricht. Auf diese Weise ließ sich die Fähigkeit zum Kritischen Denken in Anlehnung an das Triadische Modell zugleich grafisch darstellen und führte so die konzeptionelle Definition und die empirische Operationalisierung zusammen (Abbildung 34).

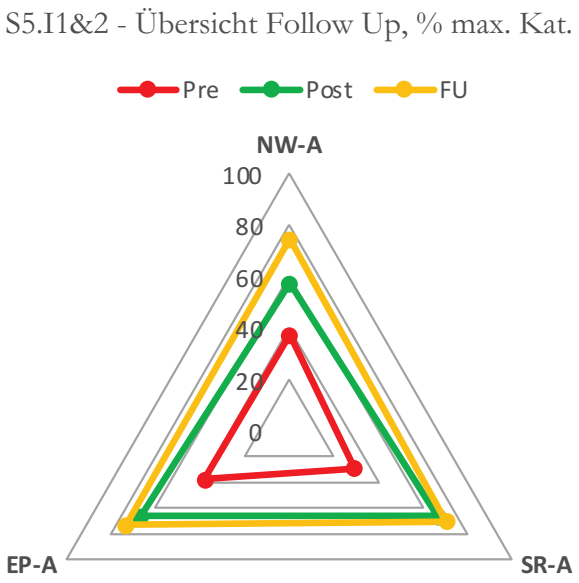


Abbildung 34: Netzdiagramm der Ergebnisse der Studie S5.I1&2 + Follow-Up

Zur Prüfung der Funktionalität und Tragfähigkeit der Definition sowie des Instruments wurden die Erhebungen der Studien S3 und S5 im Pre-Post-Follow-Up-Design konzipiert und vor Beginn sowie nach Abschluss einer eigens dafür entwickelten Vermittlungseinheit durchgeführt. Diese fokussierte sich auf die explizite Thematisierung Kritischen Denkens und unterstützte die strukturierte Auseinandersetzung mit gesellschaftlich relevanten Bewertungskontexten mit naturwissenschaftlichem Bezug (SSI) anhand der selbstkonzipierten Methode der *Systemisch-reflexiven Stellungnahme* (SrS).

Die intra- und intergruppalen Vergleiche der Ergebnisse zeigten eine hohe Stabilität des entwickelten Instruments über verschiedene Stichproben hinweg und stützen die Annahme einer hohen internen und externen Validität. Darüber hinaus wurde die Intersubjektivität durch die Ermittlung der Interkoderreliabilität in mehreren Studien abgesichert, die durchweg eine hohe Übereinstimmung erzielte und auf eine hohe Trennschärfe der Kategorien hinwies. Obwohl zur Erhebung der Bewertungsperspektiven weiterhin konkrete Bewertungsanlässe erforderlich waren, erwies sich das Instrument überdies als kontextunabhängig, was insbesondere in Studie S4 gezielt untersucht wurde. Die hohe interne und externe Validität sowie die Kontextunabhängigkeit des Instruments deuten damit auf eine vielseitige Einsetzbarkeit in unterschiedlichen Stichproben und SSI-bezogenen Kontexten hin.

Die Ergebnisse des Instruments ließen in den initialen Pre-Test-Erhebungen über alle Stichproben hinweg vorrangig intuitive sowie wenig reflektierte und begründete Bewertungsansätze erkennen, wodurch die Übertragbarkeit der Annahmen der

Dual-Process Theory gestützt wurde. Dies äußerte sich in der geringen Anzahl der bei der Bewertung von SSI herangezogenen Perspektiven, sowohl auf selbstreflexiver als auch epistemologischer und fachlicher Ebene. Zusätzlich wurde deutlich, dass relevante Begrifflichkeiten, wie Kritisches Denken oder Selbstreflexion, überwiegend subjektiv und naiv definiert vorlagen und zugleich vermittelnde selbstreflexive Prozesse in der Selbsteinschätzung der Teilnehmenden nur mäßig bis schwache Berücksichtigung fanden. Die Vergleiche zwischen den Studierenden des Bachelor- und Masterstudiums Lehramt Chemie zeigten keine signifikanten Unterschiede in den Ausgangsvoraussetzungen der Gruppen, was darauf hindeutet, dass die fachliche Studienprogression nicht zwangsläufig mit einer verbesserten Bewertungsfähigkeit einhergeht.

Im Pre-Post-Follow-Up-Vergleich wurden hingegen signifikante und nachhaltige Erweiterungen der Bewertungsperspektiven nach der Durchführung der Vermittlungseinheit und der expliziten Thematisierung Kritischen Denkens deutlich. Sowohl auf fachlicher als auch epistemologischer und selbstreflexiver Ebene nahm die Anzahl zur Bewertung hinzugezogener Aspekte qualitativ erkennbar und quantitativ signifikant zu (siehe exemplarisch Abbildung 34). Auch die Begriffsverständnisse verbesserten sich erheblich und lagen in den Post-Tests weitestgehend auf einem informierten Niveau vor. Zusätzlich zur Fremdeinschätzung verbesserte sich die Selbsteinschätzung bezüglich der Selbstreflexion dahingehend, dass sowohl die selbst erstellen Itemblöcke als auch die etablierten Testinstrumente signifikant höhere Zustimmungen erhielten und erfragten Einflussfaktoren eine verstärkte Priorität zugewiesen wurde. Dies lässt den Schluss zu, dass das Instrument das Konstrukt des Kritischen Denkens im Sinne einer multiperspektivischen Denkpraxis, welche sowohl die fachliche, epistemologische als auch selbstreflexive Dimension berücksichtigt, tatsächlich misst. Dadurch wird die Tragfähigkeit und Funktionalität der entwickelten Definition bestätigt und eine Passung zwischen der konzeptionellen Definition und der empirischen Operationalisierung hergestellt. Ferner lässt sich ableiten, dass Kritisches Denken im Sinne der erarbeiteten Definition grundsätzlich durch eine explizite Thematisierung gefördert werden kann.

Trotz der Einschränkungen der durchgeführten Untersuchungen, wie der verhältnismäßig kleinen Stichproben oder dem subjektiven Anteil in der Auswertung der qualitativen Daten, zeigen die zentralen Erkenntnisse, dass eine Operationalisierung Kritischen Denkens durch eine prozessorientierte sowie fachwissen- und kontextunabhängige Analyse der Bewertungsperspektiven in variablen gesellschaftlichen Kontex-

ten mit naturwissenschaftlichem Bezug grundsätzlich realisierbar ist. Die Verknüpfung von Fremd- und Selbsteinschätzung sowie die Implementation etablierter Instrumente ermöglichen darüber hinaus eine tiefergehende Untersuchung der wahrgenommenen Tendenz zur Selbstreflexion und eine Gegenüberstellung beider Perspektiven.

Diese Operationalisierbarkeit stärkt zusammenfassend die Tragfähigkeit und Funktionalität der entwickelten Definition Kritischen Denkens und eröffnet konkrete Perspektiven für die Weiterentwicklung von Fördermöglichkeiten zur langfristigen Etablierung Kritischen Denkens im naturwissenschaftlichen Unterricht.

6.3 Chemiedidaktische Implikationen und Ausblick

Die Ergebnisse dieser Dissertation unterstützten die Annahme, dass Wissen nicht obligat zur Positionierung in komplexen Bewertungsprozessen mit naturwissenschaftlichem Bezug herangezogen wird, sondern viele zusätzliche Faktoren auf die Anwendung von Wissen Einfluss nehmen. Rekurrierend auf das Ziel der Bildung im Allgemeinen und der naturwissenschaftlichen Grundbildung im Besonderen, Wissen in Handeln überführen und zur Orientierung in der individuellen und gesellschaftlichen Lebensrealität nutzen zu können, zeigt dies die Notwendigkeit auf, bestehende Konzepte zu ergänzen und zu erweitern. Die hier entwickelte und operationalisierte Definition Kritischen Denkens bietet konkrete Ansätze für die Verstärkung der chemiedidaktischen Vermittlung durch metakognitive Fähigkeiten, wie der Selbstreflexion. Diese metakognitiven Aspekte mögen auf den ersten Blick nicht unmittelbar mit den Naturwissenschaften in Verbindung stehen, spielen jedoch insbesondere außerhalb angeleiteter Lernumgebungen eine zentrale Rolle.

Da in der Lebensrealität häufig explizite Stimuli zur Bewertung und Reflexionen fehlen, gewinnt die eigenständige Selbstreflexion zunehmend an Bedeutung, um das im schulischen Kontext erworbene Anschlusswissen gezielt anwenden und erweitern zu können. Kritisches Denken weist dabei einen stark systemischen Charakter auf, weshalb sich das Fachwissen, das epistemologische Verständnis und die Selbstreflexion gegenseitig bedingen und auch entsprechend vermittelt werden sollten. Neben Phasen der inhaltlichen Grundlegung bedarf es daher der kontextualisierten Förderung eines Verständnisses für naturwissenschaftliche Arbeits- und Denkweisen sowie einer Integration der Selbstreflexion. Die Vermittlung Kritischen Denkens sollte jedoch prozessorientiert geschehen, indem die Breite und Tiefe der Bewertungsgrundlage und nicht allein das Ergebnis der Positionierungen in den Fokus gestellt werden.

Insbesondere, aber nicht ausschließlich, bei der Thematisierung gesellschaftlich relevanter Problemstellungen mit naturwissenschaftlichem Bezug (SSI) sowie naturwissenschaftlicher Arbeits- und Denkweisen (NOS) zeigt sich langfristig das Potenzial, Kritisches Denken in die Vermittlung einzubinden. Dies sollte, analog zur Vermittlung von NOS, explizit und kontextualisiert erfolgen, jedoch nicht additiv, sondern integrativ. Nur so kann dem systemischen Charakter des Konstrukts Rechnung getragen, eine relevante Übertragungsleistung frühzeitig angebahnt und auch die zeitliche Integration in Seminarkonzepte oder, im schulischen Kontext, Unterrichtseinheiten ermöglicht werden.

Die Ergebnisse der Studien sowie die entwickelte und erprobte Methode der Systemisch-reflexiven Stellungnahme (SrS) eröffnen in diesem Zusammenhang Perspektiven zur expliziten und angeleiteten Integration Kritischen Denkens in bestehende Lehr-Lern-Prozesse.

Im Hinblick auf die Chemielehrer*innenbildung wird sowohl die Notwendigkeit deutlich, Kritisches Denken explizit zu thematisieren, als auch die Möglichkeit, dies gewinnbringend zu realisieren. Chemielehrende nehmen hierbei als Bildungsmultiplikator*innen eine wichtige Rolle ein, da eine langfristige Verankerung Kritischen Denkens im schulischen Kontext nur gelingen kann, wenn die Lehrkräfte selbst über ein fundiertes Begriffsverständnis verfügen und in der Lage sind, multiperspektivisch zu denken. Vor dem Hintergrund dieser Verantwortung sollte Kritisches Denken zunächst explizit in die Lehrkräftebildung integriert werden, bevor eine nachhaltige Übertragung in den schulischen Unterricht gelingen kann. Die Ergebnisse der Studien deuten eine hohe Praktikabilität und Umsetzbarkeit an, da bereits kurze Seminareinheiten einen signifikanten Effekt auf die Bewertungs- und Denkstrategien der Studierenden zeigten. Über die Vermittlungsverantwortung hinaus nimmt Kritisches Denken in der Lehrer*innenbildung eine Schlüsselrolle zur kontinuierlichen professionellen Weiterentwicklung ein. Es befähigt diese, nicht nur auf fachlicher, sondern auch auf didaktischer Ebene, den eigenen Unterricht systematisch zu analysieren, zu optimieren und sowohl die zu vermittelnden Inhalte als auch die eigene Rolle und Person zu reflektieren, um damit langfristig zur Stärkung der naturwissenschaftlichen Bildung beizutragen.

Auf schulischer Ebene wohnt Kritischem Denken das Potenzial inne, die Lernenden für konkrete lebensweltliche Anforderungssituationen vorzubereiten und eine Verbindung zwischen dem Fachwissen und der eigenen Person aufzuzeigen. Damit unterstützt es das Ziel der naturwissenschaftlichen Grundbildung und legt die Grundlage zur gesellschaftlichen Teilhabe. Insbesondere im Hinblick auf die Digitalisierung

und die zunehmende Berücksichtigung dieser auch in den Bildungsstandards des Faches Chemie bietet eine Förderung Kritischen Denkens ein wichtiges Fundament für die in diesem Kontext tatsächlich kritische Auseinandersetzung mit digitalen Medien. Dies umfasst nicht nur die Nutzung digitaler Medien, sondern insbesondere die Rezeption von Informationen, den Umgang mit Künstlicher Intelligenz, die Unterscheidung zwischen wissenschaftlichen und nicht-wissenschaftlichen Inhalten, die Bewertung von Quellen sowie die Berücksichtigung damit verbundener interner und externer Einflussfaktoren auf den eigenen Entscheidungsprozess.

Als langfristige Zielsetzung erscheint eine klare und explizite Verankerung der Begrifflichkeit des Kritischen Denkens in die Lehrpläne erstrebenswert, um eine gezielte und nachhaltige Förderung zu ermöglichen. Eine fundierte und operationalisierbare Definition bildet dabei die Voraussetzung, um das oft subjektiv definierte und unscharf verwendete Wort *kritisch* inhaltlich zu präzisieren und zu einem systematisch strukturierten, praktisch anwendbaren und einheitlich definierten Terminus weiterzuentwickeln.

Die angestrebte Förderung sollte dabei kontinuierlich über mehrere Jahrgangsstufen hinweg angelegt sein und im Sinne eines Helikalcurriculums erfolgen, in dem die Elemente Kritischen Denkens schrittweise miteinander verknüpft werden und die inhaltliche Komplexität mit zunehmender Jahrgangsstufe systematisch gesteigert wird. Dieser Prozess könnte etwa von der Formulierung einfacher Fragen zu naturwissenschaftlichen Themen in den unteren Jahrgangsstufen der Sekundarstufe I bis hin zur Entwicklung komplexer, multiperspektivischer Bewertungen herausfordernder SSI in der Oberstufe reichen. Eine derartig übergreifende Förderung Kritischen Denkens würde nicht nur die praktische Umsetzung erleichtern, indem es kontinuierlich in den Lehr-Lern-Prozess eingebettet wird, sondern auch die Wahrscheinlichkeit erhöhen, dass Kritisches Denken über unterschiedliche Kontexte hinweg gefestigt und nachhaltig in die individuelle Lebensrealität der Lernenden überführt wird.

Über Wissen zu verfügen, bedeutet nicht, sich auch auf dessen Grundlage orientieren zu können. Um sich aber orientieren zu können, muss der Mensch über Wissen verfügen. Hieraus wird deutlich, dass Verfügungswissen und Orientierungswissen nicht unabhängig voneinander sind, sondern einander bedingen. Zur Überführung des Wissens von einem Verfügungswissen auf die Ebene des Orientierungswissens ist folglich nicht primär eine inhaltliche Erweiterung ausschlaggebend, sondern vielmehr die Befähigung der Lernenden, Wissen einzuordnen, zu bewerten, zu strukturieren und gezielt zu erweitern. Diese erweiterte Perspektive setzt darüber hinaus

metakognitive Fähigkeiten, wie die Selbstreflexion, voraus, um die Einflussfaktoren, die den Umgang mit Wissen prägen, zu erkennen und eine reflektierte Urteilsbildung zu ermöglichen.

Abschließend betrachtet leistet Kritisches Denken einen zentralen Beitrag zur Erreichung des Bildungsziels, Wissen in Handeln zu überführen und nachhaltige Kompetenzen zu fördern. Besonders in der heutigen Zeit, in der Wissen aufgrund des enormen Wissensumfangs nicht mehr allumfassend vermittelbar ist, Informationen jedoch durch die Digitalisierung nahezu allgegenwärtig sind, muss Bildung einerseits ein grundlegendes Anschlusswissen bereitstellen, welches eine Grundlage für eine zukünftige Wissensaneignung bildet, andererseits die Fähigkeiten schulen, Informationen bewerten, neues Wissen integrieren und Einflüsse erkennen zu können. Nur wenn Bildung diese Anforderungen systemisch adressiert, kann sie eine tatsächliche Orientierung in der Lebensrealität sowie im komplexen Gefüge gesellschaftlicher und wissenschaftlicher Diskurse ermöglichen und zur gesellschaftlichen Teilhabe befähigen.

6.4 Ausblick auf zukünftige Forschungen

Trotz der weitreichenden Implikation Kritischen Denkens und der vielversprechenden Ergebnisse des Projekts besteht weiterhin Forschungsbedarf, um Kritisches Denken langfristig sowohl in die Chemielehrer*innenbildung als auch die Schule zu integrieren. Zwar weisen die Ergebnisse darauf hin, dass eine Erfassung Kritischen Denkens über die Analyse multiperspektivischer Bewertungsprozesse umgesetzt werden kann und auch eine Förderung grundsätzlich möglich ist, dennoch bleibt die Übertragbarkeit der Befunde auf andere Kontexte und Zielgruppen eine offene Fragestellung, der in zukünftigen Studien weiter nachgegangen werden sollte.

Insbesondere eine Ausweitung auf weitere Stichproben zur Überprüfung der Generalisierbarkeit der Effekte lässt sich in diesem Kontext als sinnvoll erachten. Ergänzend wäre eine langfristige Begleitforschung wünschenswert, um die erzielten Lerneffekte und Perspektiverweiterungen weiter auf ihre Stabilität und Nachhaltigkeit der Förderansätze zu untersuchen. Darüber hinaus erscheint eine weitere systematische Erprobung der entwickelten Instrumente und Methoden in der Lehrkräftebildung sinnvoll, um deren praktische Anwendbarkeit weiter zu validieren und Optimierungspotenziale zu identifizieren. In diesem Rahmen könnten zur Validierung der Ergebnisse die Vermittlungseinheiten von unterschiedlichen Lehrenden gestaltet oder die Ergebnisse mit Kontrollgruppen verglichen werden.

Auch die Übertragbarkeit auf die Schule bleibt eine Herausforderung, da die differenten kognitiven Voraussetzungen aufgrund des entwicklungsbedingten geringeren Abstraktionsniveaus eine noch unbekannte Variable darstellen. Zukünftige Studien sollten daher gezielt die Förderung Kritischen Denkens in verschiedenen Altersstufen und Lernumgebungen adressieren und potenzielle Adaptionen oder Differenzierungsmöglichkeiten untersuchen, um zielgruppenentsprechende didaktische Konzepte zu entwickeln.

Letztlich bleibt die zentrale Herausforderung der zukünftigen Forschung nicht nur die weitere methodische Verfeinerung der Erhebung und Förderung Kritischen Denkens, sondern auch die nachhaltige Verankerung dessen als integraler Bestandteil naturwissenschaftlicher Bildung. Damit verlagert sich der Fokus weg von der Frage, *ob* Kritisches Denken definiert und operationalisiert werden kann, und hin zu der Frage, *wie* es in die Bildungssysteme implementiert werden muss, um einen nachhaltigen Beitrag zu leisten. In einer Welt, in der Informationen allgegenwärtig sind, aber deren Bewertung zunehmend anspruchsvoller wird, entscheidet sich die Qualität wissenschaftlicher und gesellschaftlicher Diskurse nicht mehr allein durch die Verfügbarkeit von Wissen, sondern durch die Fähigkeit, dieses reflektiert nutzen zu können. Kritisches Denken ist damit keine fakultative Kompetenz, sondern die Grundlage einer mündigen und zukunftsfähigen Gesellschaft, und es ist die Aufgabe der Naturwissenschaftsdidaktik, die notwendigen Konzepte bereitzustellen, um diese Grundlage tragfähig zu machen.

7 Zusammenfassung

Diese Arbeit widmet sich der Herausforderung, das komplexe Konstrukt des Kritischen Denkens zu definieren und zu operationalisieren, um einen Beitrag zu seiner langfristigen Integration und Förderung in der Chemielehrer*innenbildung sowie im schulischen Kontext zu leisten.

Zur Einordnung der Relevanz Kritischen Denkens beleuchtet der Theorieteil der Arbeit (vgl. Kap. 2) die grundsätzliche Zielsetzung von Bildung im Allgemeinen und der naturwissenschaftlichen Grundbildung im Besonderen, Wissen nicht nur zu vermitteln, sondern auch zu dessen Anwendung zu befähigen. In diesem Zusammenhang lassen sich zwei Arten von Wissen unterscheiden: Das Verfügungswissen, welches dazu befähigt, Dinge zu tun, und das Orientierungswissen, welches ermöglicht, komplexe moralische Entscheidungen zu treffen und sich in der Lebenswelt zu orientieren (Mohr, 1989) (vgl. Kap. 2.1.1). Im Rahmen der Bestrebungen der Chemiedidaktik, Kompetenzen zu fördern und ein anschlussfähiges Orientierungswissen zu vermitteln, wird im Weiteren die Bedeutung des Verständnisses naturwissenschaftlicher Arbeits- und Denkweisen im Sinne von Nature of Science (NOS) hervorgehoben und der Mensch als relevantes Element bei der Genese naturwissenschaftlicher Erkenntnisse wahrgenommen (Vgl. Kap. 2.1.2.).

Das menschliche Element nimmt jedoch nicht nur bei der Genese naturwissenschaftlicher Erkenntnisse eine Schlüsselrolle ein, sondern beeinflusst ebenfalls maßgeblich die Rezeption naturwissenschaftlicher Informationen. Besonders in der heutigen Zeit der Digitalisierung sind durch die weitreichende und unkontrollierte Verbreitung von Informationen im Internet zusätzliche Herausforderungen mit der Bewertung von Informationen verbunden. Dabei obliegt die Einschätzung der Glaubwürdigkeit den Rezipierenden selbst, welche überdies von kognitiven und sozialen Filtern beeinflusst werden (vgl. Kap. 2.1.3). Die damit verbundene Komplexität wird insbesondere bei der Konfrontation mit Fehlinformationen und nicht-wissenschaftlichen Inhalten deutlich, die, wie aktuelle Beispiele, etwa die COVID-19-Pandemie, zeigen, einen Einfluss auf die individuelle und gesellschaftliche Lebensrealität nehmen können (vgl. Kap. 2.1.4).

Die Rezeption von Informationen und ihre Einordnung werden in großem Maße von intuitiv-affektiven Faktoren geprägt, welche vor allem bei komplexen Fragestellungen Verzerrungen und Fehleinschätzungen begünstigen können. Die daraus re-

sultierende Knowledge-Action-Gap verdeutlicht, dass vorhandenes Wissen nicht obligat zu adäquaten Handlungen führt, und lässt sich an aktuellen Beispielen, wie dem Klimawandel,

veranschaulichen. Besonders bei komplexen gesellschaftlich relevanten Themen mit naturwissenschaftlichem Bezug, welche auch unter dem Begriff Socioscientific Issues (SSI) gefasst werden und sich vor allem durch ihre Uneindeutigkeit und kontroverse Natur auszeichnen, weisen bestehende Forschungen darauf hin, dass affektive Faktoren, wie persönliche Werte und Glaubenssätze, das soziale Umfeld oder die kognitiven Herausforderungen im Umgang mit inhaltlicher Komplexität, einen maßgeblichen Einfluss auf die Bewertung nehmen. Um dennoch der Zielsetzung der naturwissenschaftlichen Grundbildung zu entsprechen, Wissen in Handeln nutzbar zu machen, nimmt in Konsequenz die Forderung zu, Kritisches Denken zusätzlich zur naturwissenschaftlichen Grundbildung und einem Verständnis naturwissenschaftlicher Arbeits- und Denkweisen zu fördern.

Trotz dessen Benennung als Schlüsselkompetenz für lebenslanges Lernen und der sowohl expliziten als auch impliziten Thematisierung in nationalen und internationalen Bildungsdokumenten, verbleibt Kritisches Denken jedoch ambig definiert und infolgedessen schwer adressierbar, was die Notwendigkeit einer klaren und operationalisierbaren Definition unterstreicht. Infolgedessen wird in Kapitel 2.2 ein Einblick in bestehende Definitionen gegeben. Dabei werden nicht nur naturwissenschaftsdidaktische Ansätze, sondern auch bildungswissenschaftliche, psychologische, philosophische und historische Perspektiven in den Blick genommen. Darüber hinaus werden auch interdisziplinäre kognitionspsychologische Erkenntnisse hinzugezogen, um ein grundlegendes Verständnis für menschliche Denk- und Bewertungsprozesse zu schaffen. Von besonderer Bedeutung ist in diesem Rahmen die Dual-Process Theory, welche das Denken in das intuitive und affektiv geprägte System 1 sowie das langsamere, dafür aber rationalere System 2 differenziert (vgl. Kap. 2.3).

Unter Berücksichtigung der bestehenden Definitionsansätze und kognitionspsychologischer Erkenntnisse wird im Verlauf von Kapitel 3 eine Neudefinition Kritischen Denkens erarbeitet. Diese beschreibt Kritisches Denken als systematischen und begründeten Entscheidungsprozess, welcher die Evaluation und Anwendung des eigenen Wissens und der eigenen Fähigkeiten auf kognitiver Ebene sowie die Berücksichtigung der eigenen Emotionen, Werte, Einstellungen und Motivationen auf affektiver Ebene umfasst. In diesem Kontext steht die metakognitive Fähigkeit der Selbstreflexion im besonderen Fokus, da diese es ermöglicht, interne und externe

Einflüsse der affektiven Ebene bewusst zu reflektieren und aktiv in den Bewertungsprozess einzubeziehen (vgl. Kap. 3.3). Darauf aufbauend findet eine Abgrenzung zu bestehenden Konzepten sowie eine konkrete Übertragung auf die Chemiedidaktik statt, wobei die Relevanz Kritischen Denkens für den Bildungsprozess aufgezeigt und gezielt anhand eines entwickelten Modells konzeptualisiert wird (vgl. Kap. 3.5). Dieses Triadische Modell Kritischen Denkens charakterisiert das Konstrukt als multiperspektivische Denkpraxis, die systemisch sowohl fachliche als auch epistemologische und selbstreflexive Dimensionen in sich vereint.

Rekurrierend auf diese Definition wird im Weiteren die Operationalisierung des Konstruktes vorgenommen und die Ausschärfung des Instrumentariums über fünf iterativ aufeinander aufbauende Studien in chemiedidaktischen Seminaren der Universität zu Köln beschrieben. Dabei werden die Denk- und Bewertungsperspektiven angehender Chemielehrkräfte in konkreten gesellschaftlich relevanten Kontexten mit naturwissenschaftlichem Bezug sowohl qualitativ untersucht als auch die Ergebnisse quantitativ abgesichert. Um dem multiperspektivischen Charakter Kritischen Denkens gerecht zu werden, ist hierbei vorrangig von Bedeutung, wie viele verschiedene Aspekte in den Bewertungsprozessen berücksichtigt werden.

Der methodische Rahmen der empirischen Untersuchungen sowie die Stichproben und verwendeten Kontexte werden eingehend in Kapitel 4 beschrieben. Daran angrenzend führt das Kapitel 5 durch den iterativen Aufbau der empirischen Studien S1 bis S5, in denen insgesamt 51 Interviews und 143 Fragebögen untersucht werden. Das finale Instrument analysiert die verschiedenen Bewertungsperspektiven der Teilnehmenden innerhalb von Interviews, wobei der Bewertungsprozess als solcher und nicht nur die finale Positionierung im Fokus des Erkenntnisinteresses steht. Ferner wird den Befragten die Möglichkeit eingeräumt, für sie relevante Fragen zur Bewertung zu formulieren, um die Denkperspektiven und damit das Potenzial zur Wissensaneignung zu erfassen, anstatt primär das vorhandene Fachwissen abzubilden. Die geführten Interviews zur Fremdeinschätzung werden überdies durch Fragebogenerhebungen zur Selbsteinschätzung der Teilnehmenden hinsichtlich ihrer Selbstreflexion ergänzt, da diese eine zentrale Dimension neben den fachlichen und epistemologischen Perspektiven darstellt.

Die zentrale Auswertungsmethode bildet die deduktiv-induktive Kategorienbildung nach Kuckartz (2016), welche insbesondere bei der Interkodierung durch Elemente von Mayring (2010) ergänzt wird. Zur Sicherung der Qualität der Ergebnisse wird das Instrument in verschiedenen Stichproben und mit unterschiedlichen Kontexten

erprobt und sowohl eine qualitative Auswertung als auch quantitative Absicherung vorgenommen. Ferner wird die Interkoderreliabilität als Maß der Intersubjektivität ermittelt.

Nach seiner Ausschärfung wird zur Überprüfung der Funktionalität des Instruments eine selbstkonzipierte Vermittlungseinheit zur expliziten Förderung Kritischen Denkens erprobt und in den Studien S3 und S5 (vgl. Kap. 5.3 und 5.5) in einen Pre-Post-Test-Ansatz integriert, welcher zusätzlich durch Follow-Up-Erhebungen abgesichert wird.

Die Ergebnisse der Studien verdeutlichen, dass in den initialen Erhebungen die Positionierungen der Teilnehmenden bei der Bewertung naturwissenschaftlicher Kontexte überwiegend einseitig und wenig begründet vorgenommen wurden. Dies äußerte sich vornehmlich in der geringen Anzahl unterschiedlicher Perspektiven auf fachlicher, epistemologischer oder selbstreflexiver Ebene, die in dem Prozess berücksichtigt oder für relevant erachtet wurden (vgl. Kap. 5.7). Überdies lag das Begriffsverständnis vom Kritischen Denken stark subjektiv geladen und wenig konkret vor, während die Tendenz zur Selbstreflexion nur mäßig bis schwach ausgeprägt war. Dabei waren nahezu keine Unterschiede zwischen den Stichproben der Bachelor- und Masterstudierenden, sowohl hinsichtlich der Bewertungsperspektiven als auch in Bezug auf das Verständnis und die Tendenz, selbstreflexives Verhalten zu zeigen, erkennbar. Dies legt den Schluss nahe, dass die Studienprogression nicht zwangsläufig zu einem multiperspektivischen Denkansatz in Bewertungsprozessen führt, und deutet damit auf die Übertragbarkeit der Annahmen der Dual-Process Theory auf die Stichproben hin.

Im Gegensatz dazu zeugen die Ergebnisse der Post- und Follow-Up-Tests von einer signifikanten Erweiterung der Einbindung fachlicher, epistemologischer und selbstreflexiver Perspektiven im Bewertungsprozess. Zudem wird ein differenzierteres und weniger subjektiv geprägtes Begriffsverständnis Kritischen Denkens sichtbar, das mit einer deutlichen Zunahme der Fähigkeit zur Selbstreflexion einhergeht. Auch interne und externe Einflussfaktoren werden verstärkt in die Entscheidungsfindung einbezogen. Diese Veränderungen sind sowohl qualitativ beobachtbar als auch quantitativ messbar und legen damit die Funktionalität des Instruments und eine nachhaltige Förderung multiperspektivischen Denkens nahe.

Zur Absicherung der Ergebnisse und Überprüfung der Stabilität des Instruments werden darüber hinaus Gruppenvergleiche zwischen den Stichproben vorgenom-

men und die Kodierungen durch einen Interkoder überprüft. Die inter- und intragruppalen Vergleiche bestätigen eine hohe Stabilität des entwickelten Instruments über verschiedene Stichproben hinweg und stützen die Annahme einer hohen internen und externen Validität. Zusätzlich weisen die vorgenommenen Kodierungen große Übereinstimmungen zwischen Koder und Interkoder auf, was sich in einer sehr guten Interkoderreliabilität widerspiegelt (vgl. Kap. 5.6.).

Zusammenfassend verdeutlichen die Studienergebnisse, dass Kritisches Denken im Sinne einer multiperspektivischen Denkpraxis nicht zwangsläufig mit einer inhaltlichen Studienprogression einhergeht, sondern gezielt gefördert werden muss (vgl. Kap. 6). Gleichzeitig zeigt sich, dass eine Definition und Operationalisierung Kritischen Denkens über die Analyse multiperspektivischer Bewertungsprozesse grundsätzlich möglich ist und eine explizite Thematisierung signifikante Verbesserungen erzielt. Fachwissen im Sinne eines Verfügungswissens reicht allein nicht aus, sondern es bedarf vielmehr der Fähigkeit, fachliche, epistemologische und selbstreflexive Aspekte systematisch in die Entscheidungsprozesse zu integrieren, um sich auf Grundlage dessen orientieren zu können. Die hier entwickelte Definition, die durchgeführte Operationalisierung sowie die erprobten Vermittlungseinheiten eröffnen konkrete Perspektiven für weiterführende Forschungen, um Kritisches Denken nachhaltig curricular zu verankern. Damit verlagert sich der Schwerpunk zukünftiger Forschungen von der Frage, *ob* Kritisches Denken in die Bildung integriert werden kann, hin zur Frage, *wie* es als grundlegende Denkpraxis nachhaltig verankert werden muss.

Literaturverzeichnis

Abd-El-Khalick, F. (2006). Over And Over And Over Again: College Students' Views of Nature of Science. In L. B. Flick & N. G. Lederman (Hrsg.), *Scientific Inquiry and Nature of Science: Implications for Teaching, Learning, and Teacher Education* (Bd. 25, S. 389–425). Springer. https://doi.org/10.1007/978-1-4020-5814-1

Abd-El-Khalick, F., Bell, R. L., & Lederman, N. G. (1998). The nature of science and instructional practice: Making the unnatural natural. *Science Education*, *82*(4), 417–436. https://doi.org/10.1002/(SICI)1098-237X(199807)82:4<417::AID-SCE1>3.0.CO;2-E

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for Teaching Students to Think Critically. *Review of Educational Research*, *85*(2), 275–314. https://doi.org/10.3102/0034654314551063

Achieve, Inc. (2013). *Next Generation Science Standards*.

Adler, J. E. (2008). Philosophical Foundations. In J. E. Adler & L. J. Rips (Hrsg.), Reasoning: Studies of Human Inference and its Foundations (S. 1–34). Cambridge University Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I., & Fishbein, M. (1975). A Bayesian analysis of attribution processes. *Psychological Bulletin*, *82*(2), 261–277. https://doi.org/10.1037/h0076477

Albrecht, L. B. (2021). *Die Bedeutung naturwissenschaftlicher Arbeits- und Denkweisen für die Entwicklung und Förderung kritischen Denkens in der Chemielehrer*innenbildung* [Masterarbeit]. Universität zu Köln.

Ali, M. M. (2016). Are We Asking the Same Questions in Different Contexts: Translation Techniques in Cross-Culture Studies in Science Education? *Journal of Turkish Science Education*, *13*(1), 31–44. https://doi.org/10.12973/tused.10155a

Allelein, H.-J. (2022). Kernkraftwerke. In R. Zahoransky (Hrsg.), *Energietechnik: Systeme zur konventionellen und erneuerbaren Energieumwandlung. Kompaktwissen für Studium und Beruf* (9., überarbeitete und ergänzte Auflage, S. 104–184). Springer Vieweg.

American Psychological Association (Hrsg.). (2020). *Publication manual of the American Psychological Association (7th ed.)*. American Psychological Association. https://doi.org/10.1037/0000165-000

Aptyka, H., & Großschedl, J. (2022). Analyzing pre-service biology teachers' intention to teach evolution using the theory of planned behavior. *Evolution: Education and Outreach*, *15*(1), 16. https://doi.org/10.1186/s12052-022-00175-1

Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. (AWMF), Leitlinienprogramm Onkologie, Deutsche Krebsgesellschaft e.V. (DKG), & Stiftung Deutsche Krebshilfe (DKH). (2023). *Konsultationsfassung S3-Leitlinie Komplementärmedizin in der Behandlung von onkologischen PatientInnen* (Leitlinie AWMF-Registernummer: 032-055OL). https://www.leitlinienprogramm-onkologie.de/fileadmin/user_upload/LL_Komplement%C3%A4rmedizin_Langversion_2.01_Konsultationsfassung.pdf

Archila, P. A., Molina, J., Danies, G., Mejía, A.-M. T., & Restrepo, S. (2021). Providing Undergraduates with Opportunities to Explicitly Reflect on How News Articles Promote the Public (Mis)understanding of Science. *Science & Education*, *30*(2), 267–291. https://doi.org/10.1007/s11191-020-00175-x

Arnold, R. D., & Wade, J. P. (2015). A Definition of Systems Thinking: A Systems Approach. *Procedia Computer Science*, *44*, 669–678. https://doi.org/10.1016/j.procs.2015.03.050

BAH. (2023). *Apothekenumsatz mit rezeptfreien homöopathischen und pflanzlichen Arzneimitteln in Deutschland in den Jahren 2020 bis 2022*. https://de.statista.com/statistik/daten/studie/238427/umfrage/apothekenumsatz-mit-rezeptfreien-homoeopathischen-und-pflanzlichen-arzneimitteln/

Bailin, S. (2002). Critical Thinking and Science Education. *Science & Education*, *11*, 361–375.

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999a). Common misconceptions of critical thinking. *Journal of curriculum studies*, *31*(3), 269–283.

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999b). Conceptualizing critical thinking. *Journal of Curriculum Studies*, *31*(3), 285–302. https://doi.org/10.1080/002202799183133

Banza, C. L. N., Nawrot, T. S., Haufroid, V., Decrée, S., De Putter, T., Smolders, E., Kabyla, B. I., Luboya, O. N., Ilunga, A. N., Mutombo, A. M., & Nemery, B. (2009). High human exposure to cobalt and other metals in Katanga, a mining area of the Democratic Republic of Congo. *Environmental Research*, *109*(6), 745–752. https://doi.org/10.1016/j.envres.2009.04.012

Barr, M., & Ben-Shakhar, G. (2018). *Science, Pseudo-science, Non-sense, and Critical Thinking: Why the Differences Matter* (1. Aufl.). Routledge. https://doi.org/10.4324/9780203732892

Bausell, R. B. (2021). *The problem with science: The reproducibility crisis and what to do about it*. Oxford University Press.

Becker, U., Clarus, E., Schmidt, W., & Winter, M. (2009). *Stickoxide, Partikel und Kohlendioxid: Grenzwerte, Konflikte und Handlungsmöglichkeiten kommunaler Luftreinhaltung im Verkehrsbereich*. TU Dresden.

Beckmann, M. (2011). *Beschreibung unterschiedlicher Techniken und deren Entwicklungspotentiale zur Minderung von Stickstoffoxiden im Abgas von Abfallverbrennungsanlagen und Ersatzbrennstoff-Kraftwerken hinsichtlich Leistungsfähigkeit, Kosten und Energieverbrauch* (Umweltbundesamt, Hrsg.).

Bencze, L., Pouliot, C., Pedretti, E., Simonneaux, L., Simonneaux, J., & Zeidler, D. (2020). SAQ, SSI and STSE education: Defending and extending "science-in-context". *Cultural Studies of Science Education*, *15*(3), 825–851. https://doi.org/10.1007/s11422-019-09962-7

Berlyne, D. E. (1954). A Theory of Human Curiosity. *British Journal of Psychology. General Section*, *45*(3), 180–191. https://doi.org/10.1111/j.2044-8295.1954.tb01243.x

Bertram, A., Böttcher, C., Iyimen-Schwarz, Z., Kirschbaum, B., Osiek, D., Purr, K., & Rechenberg, J. (2014). *Fracking zur Schiefergasförderung: Eine energie- und umweltfachliche Einschätzung* (Umweltbundesamt, Hrsg.). Umweltbundesamt.

Bertram, A., Böttcher, C., Kirschbaum, B., Osiek, D., Purr, K., Rechenberg, J., & Rau, C. (2022). *Fracking zur Schiefergasförderung: Eine energie- und umweltfachliche Einschätzung* (Umweltbundesamt, Hrsg.). Umweltbundesamt.

Besley, J. C., Kramer, V. L., & Priest, S. H. (2008). Expert opinion on nanotechnology: Risks, benefits, and regulation. *Journal of Nanoparticle Research*, *10*(4), 549–558. https://doi.org/10.1007/s11051-007-9323-6

Bharadwaj, B., Baland, J. M., & Nepal, M. (2020). What makes a ban on plastic bags effective? The case of Nepal. *Environment and Development Economics*, *25*(2), 95–114. https://doi.org/10.1017/S1355770X19000329

Blancke, S., & De Smedt, J. (2013). Evolved to Be Irrational? In M. Pigliucci & M. Boudry (Hrsg.), *Philosophy of pseudoscience: Reconsidering the demarcation problem*. The University of Chicago Press.

BLK. (1997). *Gutachten zur Vorbereitung des Programms "Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts"* (Bd. 60). Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung.

Blumer, H. (1980). Der methodologische Standort des symbolischen Interaktionismus. In Arbeitsgruppe Bielefelder Soziologen (Hrsg.), *Alltagswissen, Interaktion und Gesellschaftliche Wirklichkeit* (S. 80–146). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-663-14511-0

Böhm, A. (2019). Theoretisches Codieren: Textanalyse in der Grounded Theory. In U. Flick, E. Kardorff, & I. Steinke (Hrsg.), *Qualitative Forschung: Ein Handbuch* (13. Auflage, Originalausgabe, Bd. 55628, S. 475–485). Rowohlts Enzyklopädie im Rowohlt Taschenbuch Verlag.

Bollin, E. (2022). Solartechnik. In R. Zahoransky (Hrsg.), *Energietechnik: Systeme zur konventionellen und erneuerbaren Energieumwandlung. Kompaktwissen für Studium und Beruf* (9., überarbeitete und ergänzte Auflage, S. 403–448). Springer Vieweg.

Borgeest, K. (2017). *Manipulation von Abgaswerten*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-17181-0

Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12770-0

Bosco, C. B., & Díaz, E. G. (2018). Placentophagy: A controversial trend. *International Journal of Research in Medical Sciences*, *6*(6), 1841. https://doi.org/10.18203/2320-6012.ijrms20182260

Brainard, J., & Hunter, P. R. (2020). Misinformation making a disease outbreak worse: Outcomes compared for influenza, monkeypox, and norovirus. *Simulation*, *96*(4), 365–374. https://doi.org/10.1177/0037549719885021

Brainard, J., Hunter, P. R., & Hall, I. R. (2020). An agent-based model about the effects of fake news on a norovirus outbreak. *Revue d'Épidémiologie et de Santé Publique*, *68*(2), 99–107. https://doi.org/10.1016/j.respe.2019.12.001

Brambatti Guzzo, G., & Dall'Alba, G. (2017). What Is An Ideal Critical Thinker Expected to Conclude about Anthropogenic Global Warming? *Philosophical Inquiry in Education*, *24*(3), 223–236. https://doi.org/10.7202/1070608ar

Brennan, R. L., & Prediger, D. J. (1981). Coefficient Kappa: Some Uses, Misuses, and Alternatives. *Educational and Psychological Measurement*, *41*(3), 687–699. https://doi.org/10.1177/001316448104100307

Brettschneider, F., & Müller, U. (2020). Vorhabenträger auf dem Weg zu gesellschaftlich tragfähigen Lösungen. In F. Brettschneider (Hrsg.), *Bau- und Infrastrukturprojekte: Dialogorientierte Kommunikation als Erfolgsfaktor* (S. 1–40). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-28235-6

Brislin, R. W. (1970). Back-Translation for Cross-Cultural Research. *Journal of Cross-Cultural Psychology*, *1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Brislin, R. W. (1976). Comparative research methodology: Cross-cultural studies. *International Journal of Psychology*, *11*(3), 215–229. https://doi.org/10.1080/00207597608247359

Bruder, R., Leuders, T., & Büchter, A. (2017). *Mathematikunterricht entwickeln: Bausteine für kompetenzorientiertes Unterrichten* (5. Aufl.). Cornelsen.

Bründlinger, R., Christ, D., Fechner, H., Kaltschmitt, M., Müller, J., Peharz, G., Schulz, D., & Sens, L. (2020). Photovoltaische Stromerzeugung. In M. Kaltschmitt, W. Streicher, & A. Wiese (Hrsg.), *Erneuerbare Energien: Systemtechnik, Wirtschaftlichkeit, Umweltaspekte* (6., vollständig neu überarbeitete Auflage, S. 339–460). Springer Vieweg.

Bucher, U., Holzweißig, K., & Schwarzer, M. (2024). *Künstliche Intelligenz und wissenschaftliches Arbeiten: ChatGPT & Co.: der Turbo für ein erfolgreiches Studium*. Verlag Franz Vahlen.

Bund für Umwelt und Naturschutz Deutschland, Nuclear Free Future Foundation, Rosa-Luxemburg-Stiftung, & Umweltstiftung Greenpace (Hrsg.). (2022). *Uranatlas—Daten und Fakten über den Rohstoff des Atomzeitalters* (2. Aufl.). www.bund.net/uranatlas

Bundesministerium für Familie, Senioren, Frauen und Jugend. (o. J.). *Verschwörungsmythen über die Corona-Pandemie*. Demokratie leben! Abgerufen 7. August 2024, von https://www.demokratie-leben.de/magazin/magazin-details/verschwoerungsmythen-ueber-die-corona-pandemie-44

Bundesministerium für Wirtschaft und Klimaschutz (BMWK) (Hrsg.). (2024). *Monitoringbericht zum Ausbau der erneuerbaren Energien im Strombereich nach § 98 Absatz 3 EEG und Fortschrittsbericht Windenergie an Land nach § 99a EEG - Bericht der Bundesregierung*. https://www.bmwk.de/Redaktion/DE/Publikationen/Energie/monitoringbericht-zum-ausbau-der-erneuerbaren-energien-im-strombereich-2023.html

Bundesverband der Arzneimittel-Hersteller e.V. (BAH). (2023). *Der Arzneimittelmarkt in Deutschland—Zahlen & Fakten aus 2022* [Statistik]. Bundesverband der Arzneimittel-Hersteller e.V. (BAH). https://www.pharmadeutschland.de/index.php?id=1&type=565&file=redakteur_filesystem/public/Weitere_oeffentliche_Dateien/20230522_BAH_Zahlenbroschuere_2022_WEB.pdf

Buser, G. L., Mató, S., Zhang, A. Y., Metcalf, B. J., Beall, B., & Thomas, A. R. (2017). Late-Onset Infant Group B Streptococcus Infection Associated with Maternal Consumption of Capsules Containing Dehydrated Placenta—Oregon, 2016. *Morbidity and Mortality Weekly Report*, *66*(25).

Busker, M. (2014). Entwicklung eines Fragebogens zur Untersuchung des Fachinteresses. In D. Krüger, I. Parchmann, & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 269–282). Springer Berlin Heidelberg.

Butler, T., von Schneidemesser, E., & Becker, S. (2018). *Bekämpfung der städtischen Luftverschmutzung: Stickoxide und Dieselemissionen*. IASS Policy Brief.

Bybee, R. W. (1997). *Achieving scientific literacy: From purposes to practices*. Heinemann.

Bybee, R. W. (2002). Scientific Literacy -Mythos oder Realität? In W. Gräber, P. Nentwig, T. Koballa, & R. Evans (Hrsg.), *Scientific Literacy*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-80863-9

Bybee, R. W., & Roberts, D. A. (2014). Scientific Literacy, Science Literacy, and Science Education. In N. G. Lederman & S. K. Abell, *Handbook of research on science education Vol. II*. Routledge.

Byrne, M. S., & Johnstone, A. H. (1987). Critical thinking and science education. *Studies in Higher Education*, *12*(3), 325–339. https://doi.org/10.1080/03075078712331378102

Cândido, L., Rabelo, J. N. T., Da Silva, J. L. F., & Hai, G.-Q. (2012). Quantum Monte Carlo study of small aluminum clusters Al n (n = 2 – 13). *Physical Review B*, *85*(24), 245404. https://doi.org/10.1103/PhysRevB.85.245404

Carver, C. S., & Scheier, M. F. (2001). *On the self-regulation of behavior* (1. Paperback Aufl.). Cambridge Univ. Press.

Cebesoy, U. B., & Oztekin, C. (2018). Genetics Literacy: Insights From Science Teachers' Knowledge, Attitude, and Teaching Perceptions. *International Journal of Science and Mathematics Education*, *16*(7), 1247–1268. https://doi.org/10.1007/s10763-017-9840-4

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*(5), 752–766. https://doi.org/10.1037/0022-3514.39.5.752

Chapman, G. B., & Johnson, E. J. (2002). Incorporating the Irrelevant: Anchors in Judgments of Belief and Value. In T. Gilovich, D. Griffin, & D. Kahneman (Hrsg.), *Heuristics and biases: The psychology of intuitive judgement* (S. 120–138). Cambridge University Press.

Chen, L., & Xiao, S. (2021). Perceptions, challenges and coping strategies of science teachers in teaching socioscientific issues: A systematic review. *Educational Research Review*, *32*, 100377. https://doi.org/10.1016/j.edurev.2020.100377

Christel, H. (2019). Qualitative Interviews—Ein Überblick. In U. Flick, E. Kardorff, & I. Steinke (Hrsg.), *Qualitative Forschung: Ein Handbuch* (13. Auflage, Originalausgabe, Bd. 55628, S. 349–359). Rowohlts Enzyklopädie im Rowohlt Taschenbuch Verlag.

Clauss, G., & Ebner. (1971). *Grundlagen der Statistik*. Harri Deutsch.

Cleff, T. (2015). *Deskriptive Statistik und Explorative Datenanalyse: Eine computergestützte Einführung mit Excel, SPSS und STATA*. Gabler Verlag. https://doi.org/10.1007/978-3-8349-4748-2

Clough, M. P. (2006). Learners' Responses to the Demands of Conceptual Change: Considerations for Effective Nature of Science Instruction. *Science & Education*, *15*(5), 463–494. https://doi.org/10.1007/s11191-005-4846-7

Clough, M. P. (2008). Teaching the nature of science to secondary and post-secondary students: Questions rather than tenets. *California Journal of Science Education*, *8*(2), 31–40.

Cobb, M. D., & Macoubrie, J. (2004). Public perceptions about nanotechnology: Risks, benefits and trust. *Journal of Nanoparticle Research*, *6*, 395–405.

Cohen, B. H. (2001). *Explaining psychological statistics* (2. Aufl.). J. Wiley & Sons.

Cohen, B. H. (2008). *Explaining psychological statistics* (3. Aufl.). Wiley.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment Vol 2: Standard issue scoring manual* (2. Aufl.). Cambridge Univ. Press.

Coleman, J. S. (1990). *Foundations of social theory* (3. Aufl.). Belknap Press of Harvard Univ. Press.

Collins, H. (2009). We cannot live by scepticism alone. *Nature*, *458*(7234), 30–30. https://doi.org/10.1038/458030a

Colombo, S. L., Chiarella, S. G., Lefrançois, C., Fradin, J., Raffone, A., & Simione, L. (2023). Why Knowing about Climate Change Is Not Enough to Change: A Perspective Paper on the Factors Explaining the Environmental Knowledge-Action Gap. *Sustainability*, *15*(20), 14859. https://doi.org/10.3390/su152014859

Comenius, J. A., & Overhoff, J. (Hrsg.) (2022). *Große Didaktik: Die vollständige Kunst, alle Menschen alles zu lehren* (A. Flitner, Übers.; 1. Aufl.). Klett-Cotta.

Connect Jugendmedienzentrum, Fanprojekt Fürth, & Echt Fürth. (o. J.). *Realtalk—Verschwörungsmythen in der Corona-Krise*. Echt Fürth. Abgerufen 7. August 2024, von https://echt-fuerth.de/realtalk/

Cosmides, L., & Tooby, J. (2008). Beyond Intuition and Instinct Blindness: Toward an Evolutionarily Rigorous Cognitive Science. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and its Foundations* (S. 844–865). Cambridge University Press.

Coyle, C. W., Hulse, K. E., Wisner, K. L., Driscoll, K. E., & Clark, C. T. (2015). Placentophagy: Therapeutic miracle or myth? *Archives of Women's Mental Health*, *18*(5), 673–680. https://doi.org/10.1007/s00737-015-0538-8

Creswell, J. W. (2015). *A concise introduction to mixed methods research*. Sage.

Cucherat, M., Haugh, M. C., Gooch, M., & Boissel, J.-P. (2000). Evidence of clinical efficacy of homeopathy. *European Journal of Clinical Pharmacology*, *56*(1), 27–33. https://doi.org/10.1007/s002280050716

Dachler, M. (2023). *Welternährung: Status quo und Ausblick zur globalen Ernährungslage*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-66904-4

De Neys, W. (2006a). Automatic–Heuristic and Executive–Analytic Processing during Reasoning: Chronometric and Dual-Task Considerations. *Quarterly Journal of Experimental Psychology*, *59*(6), 1070–1100. https://doi.org/10.1080/02724980543000123

De Neys, W. (2006b). Dual Processing in Reasoning: Two Systems but One Reasoner. *Psychological Science*, *17*(5), 428–433. https://doi.org/10.1111/j.1467-9280.2006.01723.x

De Sousa, R. (2008). Biology, Emotions, and Reasoning. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and its Foundations* (S. 1002–1015). Cambridge University Press.

DECHEMA (Hrsg.). (2015). *Anorganische Rohstoffe—Sicherung der Rohstoffbasis von morgen. Positionspapier des Temporären ProcessNet-Arbeitskreises „Rohstoffe und Kreislaufwirtschaft"*. DECHEMA Gesellschaft für Chemische Technik und Biotechnologie e.V.

Decker, O., Schuler, J., Yendell, A., Schließler, C., & Brähler, E. (2020). Das autoritäre Syndrom: Dimensionen und Verbreitung der Demokratie-Feindlichkeit. In O. Decker & E. Brähler (Hrsg.), *Autoritäre Dynamiken: Alte Ressentiments—Neue Radikalität: Leipziger Autoritarismus Studie 2020* (Originalausgabe). Psychosozial-Verlag.

Denzin, N. K. (2017). *The research act: A theoretical introduction to sociological methods*. Routledge, Taylor & Francis Group. https://doi.org/10.4324/9781315134543

Deutscher Bundestag, W. D. (2022). *Seltene Erden als wichtige Ressource* (Nummer WD 5-3000-003/22, S. 1-22).

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process* (Neue Aufl.). D. C. Heath.

Dewey, J., Hylla, E., & Dewey, J. (2011). *Demokratie und Erziehung: Eine Einleitung in die philosophische Pädagogik ; mit einer umfangreichen Auswahlbibliographie* (J. Oelkers, Hrsg.; 5. Aufl., [Nachdr. der Ausg. Weinheim, Beltz 1993]). Beltz.

Diemand, S. (2023). Das Geschäft mit dem Aberglauben. *Frankfurter Allgemeine Zeitung*. https://www.faz.net/aktuell/wirtschaft/schneller-schlau/das-geschaeft-mit-dem-aberglauben-esoterik-ist-ein-milliardenschwerer-markt-18946519.html

Dodge, Y. (2010). *The concise encyclopedia of statistics*. Springer.

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-41089-5

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, *65*(2), 135–150.

Dr. Dresing & Pehl GmbH. (2023). *Audiotranskription: F4x automatische Spracherkennung*. https://www.audiotranskription.de/f4x/

Dr. Dresing & Pehl GmbH. (2024). *F4 – Software für Transkription & qualitative Auswertung*. https://www.audiotranskription.de/shop/f4-software-fuer-qualitative-forschungsprojekte/

Dresing, T., & Pehl, T. (2018). *Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende* (8. Auf.). Eigenverlag.

Driver, Rosalind., Leach, John., Millar, R., & Scott, P. (1996). *Young people's images of science*. Open Univ. Press.

Dubs, R. (2002). Literacy: A great challenge for pedagogy. In *Qualitätsentwicklung im naturwissenschaftlichen Unterricht. Fachtagung am 15. Dezember 1999: Bd. Materialien zur Bildungsforschung*. GFPF; DIPF: Frankfurt am Main.

Duit, R., Pfund, H., & Jung, W. (1981). *Alltagsvorstellungen und naturwissenschaftlicher Unterricht*. Aulis-Verlag Deubner.

Duller, C. (2019). *Einführung in die Statistik mit EXCEL und SPSS: Ein anwendungsorientiertes Lehr- und Arbeitsbuch*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-59410-0

Dunning, D., Meyerowitz, J. A., & Holzberg. (2002). Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability. In T. Gilovich, D. Griffin, & D. Kahneman (Hrsg.), *Heuristics and biases: The psychology of intuitive judgement* (S. 324–333). Cambridge University Press.

Dupont, L., Jacob, S., & Philippe, H. (2024). Scientist engagement and the knowledge–action gap. *Nature Ecology & Evolution*. https://doi.org/10.1038/s41559-024-02535-0

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity*, *12*, 43–52. https://doi.org/10.1016/j.tsc.2013.12.004

Elio, R., & Pelletier, F. J. (2008). Belief Change as Propositional Update. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and its Foundations*. Cambridge University Press.

Elster, J. (1989). *The cement of society: A study of social order*. Cambridge University Press.

Elster, J. (2008). When Rationality Fails. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and its Foundations* (S. 94–113). Cambridge University Press.

Ennis, R. H. (1987). A Taxonomy of Critical Thinking Dispositions and Abilities. In J. B. Baron & R. J. Sternberg (Hrsg.) *Teaching Thinking Skills: Theory and Practice*. W.H. Freeman and Company.

Ennis, R. H. (1989). *Critical Thinking and Subject Specificity: Clarification and Needed Research*. *18*(3), 4–10.

Epstein, S. (1994). Integration of the Cognitive and Psychodynamic Unconscious. *American psychologist*, *48*(8), 709.

Erb, S., & Fichtner, S. (2023, April 28). Die gekaufte Professur—Anthroposophische Medizin an der Charité. *Die Tageszeitung*. https://taz.de/Anthroposophische-Medizin-an-der-Charite/!5930859/

Erdmann, M., & Bundesanstalt für Geowissenschaften und Rohstoffe. (2021). *Seltene Erden: Informationen zur Nachhaltigkeit*. https://www.bgr.bund.de/DE/Themen/Min_rohstoffe/Produkte/Informationen_zur_Nachhaltigkeit/seltene%20erden_verzeichnis.html

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data* (2. Aufl.). MIT Press.

Ernst, E. (2008). Chiropractic: A Critical Evaluation. *Journal of Pain and Symptom Management*, *35*(5), 544–562. https://doi.org/10.1016/j.jpainsymman.2007.07.004

Ernst, E. (2015). Ehemaliger Homöopath warnt: „Homöopathie kann lebensgefährlich sein". *Geo*, *15*(9).

Europäische Kommission. (2011). *Grundstoffmärkte und Rohstoffe: Herausforderungen und Lösungsansätze*. https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52011DC0025&from=GA

Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review*, *13*(3), 378–395. https://doi.org/10.3758/BF03193858

Evans, J. St. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, *59*(1), 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Evans, J. St. B. T. (2018). Dual-Process Theories. In L. J. Ball & V. A. Thompson (Hrsg.), *The Routledge international handbook of thinking and reasoning* (S. 151–166). Routledge.

Eyerer, P. (2020). Einführung in Polymer Engineering. In P. Eyerer & H. Schüle (Hrsg.), *Polymer Engineering 1: Einführung, Synthese, Eigenschaften* (S. 1–44). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-59837-5

Eyrer, P. (2020). Synthese (Herstellung, Erzeugung) von (petrobasierten) Kunststoffen. In P. Eyerer & H. Schüle (Hrsg.), *Polymer Engineering 1: Einführung, Synthese, Eigenschaften* (S. 45–88). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-59837-5

Facione, P. A. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. Research Findings and Recommendations*. American Philosophical Association.

Fahrner, W. R., & Ulyashin, A. (2017). Quantenmechanische Aspekte. In W. Fahrner (Hrsg.), *Nanotechnologie und Nanoprozesse* (S. 7–18). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-48908-6

Farr, A., Chervenak, F. A., McCullough, L. B., Baergen, R. N., & Grünebaum, A. (2018). Human placentophagy: A review. *American Journal of Obstetrics and Gynecology*, *218*(4), 401-410. https://doi.org/10.1016/j.ajog.2017.08.016

Fasce, A. (2017). What do we mean when we speak of pseudoscience? The development of a demarcation criterion based on the analysis of twenty-one previous attempts. *Disputatio. Philosophical Research Bulletin*, *6*(7), 459–488. https://doi.org/10.5281/zenodo.1433737

Fath, A. (2019). *Mikroplastik: Verbreitung, Vermeidung, Verwendung*. Springer Spektrum.

Faulstich-Christ, K., Lersch, R., & Moegling, K. (Hrsg.). (2010). *Kompetenzorientierung in Theorie, Forschung und Praxis: Sekundarstufen I und II* (1. Aufl.). Verlag Barbara Budrich. https://doi.org/10.2307/j.ctvdf0dmc

Feierabend, S., Rathgeb, T., Kheredmand, H., & Glöckler, S. (2021). *JIM-Studie 2021: Jugend, Information, Medien. Basisuntersuchungen zum Medienumgang 12- bis 19-Jähriger*. Medienpädagogischer Forschungsverbund Südwest (MPFS). https://www.mpfs.de/fileadmin/files/Studien/JIM/2021/JIM-Studie_2021_barrierefrei.pdf;

Feierabend, S., Rathgeb, T., Kheredmand, H., & Glöckler, S. (2022). *JIM-Studie 2022: Jugend, Information, Medien. Basisuntersuchungen zum Medienumgang 12- bis 19-Jähriger*. Medienpädagogischer Forschungsverbund Südwest (MPFS). https://www.mpfs.de/fileadmin/files/Studien/JIM/2022/JIM_2022_Web_final.pdf;

Festinger, L. (1957). *A theory of cognitive dissonance*. Row, Petersen.

Fichtner, C. (2022). Windenergie. In R. Zahoransky (Hrsg.), *Energietechnik: Systeme zur konventionellen und erneuerbaren Energieumwandlung. Kompaktwissen für Studium und Beruf* (9., überarbeitete und ergänzte Aufl., S. 449–520). Springer Vieweg.

Fixen, P. E., & Johnston, A. M. (2012). World fertilizer nutrient reserves: A view to the future. *Journal of the Science of Food and Agriculture*, *92*(5), 1001–1005. https://doi.org/10.1002/jsfa.4532

Fläming, H., Gertz, C., & Mühlhausen, T. (2023). Minderungsstrategien im Personen- und Güterverkehr. In G. P. Brasseur, D. Jacob, & S. Schuck-Zöller (Hrsg.), *Klimawandel in Deutschland: Entwicklung, Folgen, Risiken und Perspektiven* (S. 415–428). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-66696-8

Flick, U. (2011). *Triangulation: Eine Einführung* (3. aktualisierte Auflage, Bd. 12). VS Verl. für Sozialwiss. https://doi.org/10.1007/978-3-531-92864-7

Flick, U., Kardorff, E., & Steinke, I. (Hrsg.). (2019). *Qualitative Forschung: Ein Handbuch* (13. Auflage, Originalausgabe, Bd. 55628). Rowohlt Taschenbuch Verlag.

Franco, A. R., & Vieira, R. M. (2019, Juni 26). Promoting critical thinking in higher education in the context of teacher professional development. *5th International Conference on Higher Education Advances (HEAd'19)*. Fifth International Conference on Higher Education Advances. https://doi.org/10.4995/HEAD19.2019.9077

Frankish, K. (2010). Dual-Process and Dual-System Theories of Reasoning. *Philosophy Compass*, *5*(10), 914–926. https://doi.org/10.1111/j.1747-9991.2010.00330.x

Frederick, S. (2002). Automated Choice Heuristics. In T. Gilovich, D. Griffin, & D. Kahneman (Hrsg.), *Heuristics and biases: The psychology of intuitive judgement* (S. 548–558). Cambridge University Press.

Frederick, S. (2005). Cognitive reflection and decision making. *The journal of economic perspectives*, *19*(4), 25–42.

Freud, S. (2010). The interpretation of dreams (J. Strachey, Übers.; Basic Books).

Frimmel, F. H., Ewers, U., Schmitt-Jansen, M., Gordalla, B., & Altenburger, R. (2012). Toxikologische Bewertung von Fracking-Fluiden. *Wasser und Abfall*, *14*(6), 22–29.

Frische, T., Egerer, S., Matezki, S., Pickl, C., & Wogram, J. (2016). *5-Punkte-Programm für einen nachhaltigen Pflanzenschutz* (Umweltbundesamt, Hrsg.). Umweltbundesamt.

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*(1), 2–18. https://doi.org/10.1037/a0024338

Früh, W. (2017). *Inhaltsanalyse: Theorie und Praxis* (9., überarbeitete Aufl., Online-Ausgabe). UVK. https://doi.org/10.36198/9783838547350

Fuchs, C. (2022). *Verschwörungstheorien in der Pandemie: Wie über COVID-19 im Internet kommuniziert wird* (1. Aufl.). utb GmbH. https://doi.org/10.36198/9783838557960

Gasparatos, A., Doll, C. N. H., Esteban, M., Ahmed, A., & Olang, T. A. (2017). Renewable energy and biodiversity: Implications for transitioning to a Green Economy. *Renewable and Sustainable Energy Reviews*, *70*, 161–184. https://doi.org/10.1016/j.rser.2016.08.030

Geertz, C. (2019). *Dichte Beschreibung: Beiträge zum Verstehen kultureller Systeme* (B. Luchesi & R. Bindemann, Übers.; 14. Aufl.). Suhrkamp.

Ghebreyesus, T. A. (2020, Februar 15). *Rede von Dr. Tedros Adhanom Ghebreyesus, Generaldirektor der WHO*. Münchner Sicherheitskonferenz, München. https://www.who.int/director-general/speeches/detail/munich-security-conference

Ghisellini, F., & Chang, B. Y. (2018). *Behavioral Economics*. Springer International Publishing. https://doi.org/10.1007/978-3-319-75205-1

Gigerenzer, G. (1991). How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases". *European Review of Social Psychology*, *2*(1), 83–115. https://doi.org/10.1080/14792779143000033

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgement*. Cambridge University Press.

Glaser, B. G., & Strauss, A. L. (1965). Discovery of Substantive Theory: A Basic Strategy Underlying Qualitative Research. *American Behavioral Scientist*, *8*(6), 5–12. https://doi.org/10.1177/000276426500800602

Gläser, J., & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse: Als Instrumente rekonstruierender Untersuchungen* (3. Aufl.). VS Verl.

Gläser-Zikuda, M., Stephan, M., & Hofmann, F. (2022). Qualitative Auswertungsverfahren. In H. Reinders, D. Bergs-Winkels, A. Prochnow, & I. Post (Hrsg.), *Empirische Bildungsforschung: Eine elementare Einführung* (S. 237–252). Springer VS. https://doi.org/10.1007/978-3-658-27277-7

Godfrey, L. (2019). Waste Plastic, the Challenge Facing Developing Countries—Ban It, Change It, Collect It? *Recycling*, *4*(1), 3. https://doi.org/10.3390/recycling4010003

Goodship, V. (2007). Plastic Recycling. *Science Progress*, *90*(4), 245–268. https://doi.org/10.3184/003685007X228748

Gräber, W. (2002). „Scientific Literacy"—Naturwissenschaftliche Bildung in der Diskussion. In *Qualitätsentwicklung im naturwissenschaftlichen Unterricht. Fachtagung am 15. Dezember 1999: Bd. Materialien zur Bildungsforschung* (Nummer 7, S. 1–28). GFPF; DIPF: Frankfurt am Main. https://doi.org/10.25656/01:3443

Gräber, W., & Nentwig, P. (2002). Scientific Literacy—Naturwissenschaftliche Grundbildung in der Diskussion. In W. Gräber, P. Nentwig, T. Koballa, & R. Evans (Hrsg.), *Scientific Literacy*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-80863-9

Grant, A. M., Franklin, J., & Langford, P. (2002). The Self-Reflection and Insight Scale: A New Measure of Private Self-Consciousness. *Social Behavior and Personality: an international journal*, *30*(8), 821–835. https://doi.org/10.2224/sbp.2002.30.8.821

Greene, J. A., & Yu, S. B. (2016). Educating Critical Thinkers. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 45–53. https://doi.org/10.1177/2372732215622223

Greenhalgh, T. (2002). Intuition and evidence—Uneasy bedfellows? *British Journal of General Practice*, *52*(478), 395–400.

Groeben, N., & Rustemeyer, R. (1995). Inhaltsanalyse. In E. König (Hrsg.), *Bilanz qualitativer Forschung. 2: Methoden* (S. 523–554). Dt. Studien-Verl.

Groß, K. (2022). Individuelle Förderung im Chemieunterricht. In C. S. Reiners, *Chemie vermitteln: Fachdidaktische Grundlagen und Implikationen* (2. erweiterte Aufl., S. 183–204). Springer Spektrum. https://doi.org/10.1007/978-3-662-65511-5

Guillaud, A., Darbois, N., Monvoisin, R., & Pinsault, N. (2018). Reliability of diagnosis and clinical efficacy of visceral osteopathy: A systematic review. *BMC Complementary and Alternative Medicine*, *18*(1), 65. https://doi.org/10.1186/s12906-018-2098-8

Gunn, T. M., Grigg, L. M., & Pomahac, G. A. (2008). Critical Thinking in Science Education: Can Bioethical Issues and Questioning Strategies Increase Scientific Understandings? *The Journal of Educational Thought (JET)*, *42*(2), 165–183.

Gupta, V. K., Saini, C., Oberoi, M., Kalra, G., & Nasir, M. I. (2020). Semmelweis Reflex: An Age-Old Prejudice. *World Neurosurgery*, *136*, 119–125. https://doi.org/10.1016/j.wneu.2019.12.012

Habrich-Böcker, C., Kirchner, B., & Weißenberg, P. (2015). *Fracking: Die neue Produktionsgeografie* (2., aktualisierte und korr. Aufl.). Springer Gabler.

Hagenauer, G., & Gläser-Zikuda, M. (2022). Mixed Methods. In H. Reinders, D. Bergs-Winkels, A. Prochnow, & I. Post (Hrsg.), *Empirische Bildungsforschung: Eine elementare Einführung* (S. 253–268). Springer VS. https://doi.org/10.1007/978-3-658-27277-7

Haidt, J. (2008). The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. In J. E. Adler & L. J. Rips, *Reasoning: Studies of Human Inference and its Foundations*. Cambridge University Press.

Hall, H. G., McKenna, L. G., & Griffiths, D. L. (2012). Midwives' support for Complementary and Alternative Medicine: A literature review. *Women and Birth*, *25*(1), 4–12. https://doi.org/10.1016/j.wombi.2010.12.005

Haller, L., Moakes, S., Niggli, U., Riedel, J., Stolze, M., & Thompson, M. (2020). *Entwicklungsperspektiven der ökologischen Landwirtschaft in Deutschland* (Umweltbundesamt, Hrsg.). Umweltbundesamt. https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/2020-03-17_texte_32-2020_oekologische-landwirtschaft.pdf

Halpern, D. F. (1999). Teaching for Critical Thinking: Helping College Students Develop the Skills and Dispositions of a Critical Thinker. *New Directions for Teaching and Learning*, *1999*(80), 69–74. https://doi.org/10.1002/tl.8005

Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking* (5. Aufl.). Psychology Press.

Hamacher, T. (2017). Wasserstoff als strategischer Sekundärenergieträger. In J. Töpler, J. Lehmann, & E. U. von Weizsäcker (Hrsg.), *Wasserstoff und Brennstoffzelle: Technologien und Marktperspektiven* (2., aktualisierte und erweiterte Auflage, S. 1–24). Springer Vieweg. https://doi.org/10.1007/978-3-662-53360-4

Hammann, M., & Jördens, J. (2014). Offene Aufgaben codieren Marcus Hammann und Janina Jördens. In D. Krüger, I. Parchmann, & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 169–178). Springer Berlin Heidelberg.

Hanfeld, M. (2020, Mai 26). Die versuchte Vernichtung. *Frankfurter Allgemeine Zeitung*.

Hann, S., & Connock, T. (2020). *Chemical Recycling: State of Play* (Eunomia Research & Consulting Ltd, Hrsg.). https://chemtrust.org/wp-content/uploads/Chemical-Recycling-Eunomia.pdf

Hannah, W., & Thompson, P. B. (2008). Nanotechnology, risk and the environment: A review. *Journal of Environmental Monitoring*, *10*(3), 291. https://doi.org/10.1039/b718127m

Hansson, S. O. (2013). Defining Pseudoscience and Science. In M. Pigliucci & M. Boudry (Hrsg.), *Philosophy of pseudoscience: Reconsidering the demarcation problem* (S. 61–77). The University of Chicago Press.

Harmon-Jones, E., & Mills, J. (2019). An introduction to cognitive dissonance theory and an overview of current perspectives on the theory. In E. Harmon-Jones (Hrsg.), *Cognitive dissonance: Reexamining a pivotal theory in psychology* (2. Aufl., S. 3–24). American Psychological Association. https://doi.org/10.1037/0000135-001

Harrington, R., & Loffredo, D. A. (2010). Insight, Rumination, and Self-Reflection as Predictors of Well-Being. *The Journal of Psychology*, *145*(1), 39–57. https://doi.org/10.1080/00223980.2010.528072

Hartmann, N., Saad Hussein, N., Taumann, Jülch, V., & Schlegl, T. (2015). Stromerzeugung aus Windenergie. In M. Wietschel, S. Ullrich, P. Markewitz, F. Schulte, & F. Genoese (Hrsg.), *Energietechnologien der Zukunft: Erzeugung, Speicherung, Effizienz und Netze* (S. 103–122). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-07129-5

Hays, W. L., & Winkler, R. L. (1970). *Statistics* (Bd. 1). Holt, Rinehart and Winston.

Heinrich, J. (2018). Feinstaub, Stickoxide, Ozon: Darf man in Großstädten noch atmen? *Der Pneumologe*, *15*(4), 263–265. https://doi.org/10.1007/s10405-018-0192-8

Hellweg, N., Schließler, C., & Decker, O. (2020). Aberglaube, Esoterik und Verschwörungsmentalität in Zeiten der Pandemie. In O. Decker & E. Brähler (Hrsg.), *Autoritäre Dynamiken: Alte Ressentiments—Neue Radikalität: Leipziger Autoritarismus Studie 2020* (Originalausgabe). Psychosozial-Verlag.

Helmers, E. (2009). Partikelmessungen, Abgasgrenzwerte, Stickoxide, Toxikologie und Umweltzonen. *Umweltwissenschaften und Schadstoff-Forschung*, *21*(1), 118–123. https://doi.org/10.1007/s12302-008-0026-0

Hengstmann, E., & Tamminga, M. (2022). *Plastik in der Umwelt: Wo kommt es her, wo geht es hin und wie wirkt es sich aus?* Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-65864-2

Herberz, T., Barlow, C. Y., & Finkbeiner, M. (2020). Sustainability Assessment of a Single-Use Plastics Ban. *Sustainability*, *12*(9), 3746. https://doi.org/10.3390/su12093746

Hess, D. J., & Lamprou, A. (2012). Nanotechnology and the Environment. In D. Maclurcan & N. Radywyl (Hrsg.), *Nanotechnology and global sustainability* (S. 21–44). CRC Press.

Hildebrand, B. (2019). Anselm Strauss. In U. Flick, E. Kardorff, & I. Steinke (Hrsg.), *Qualitative Forschung: Ein Handbuch* (13. Auflage, Originalausgabe, Bd. 55628, S. 32–42). Rowohlts Enzyklopädie im Rowohlt Taschenbuch Verlag.

Hilton, D. J. (2008). The Social Context of Reasoning: Conversational Inference and Rational Judgment. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and its Foundations* (S. 774–806). Cambridge University Press.

Hoffrage, U., Hafenbrädel, S., & Marewski, J. N. (2018). The Fast and Frugal Heuristics Programm. In L. J. Ball & V. A. Thompson (Hrsg.), *The Routledge international handbook of thinking and reasoning* (S. 325–345). Routledge.

Holbrook, J. (2010). Education through science as a motivational innovation for science education for all. *Science Education International*, *21*(2), 80–91.

Hopewell, J., Dvorak, R., & Kosior, E. (2009). Plastics recycling: Challenges and opportunities. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1526), 2115–2126. https://doi.org/10.1098/rstb.2008.0311

Hopp, V. (2018). *Chemische Kreisläufe in der Natur: Chemie - Biologie - Energetik*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-55860-7

Horneber, M., Van Ackeren, G., Linde, K., & Rostock, M. (2008). Mistletoe therapy in oncology. *Cochrane Database of Systematic Reviews*, *2020*(11). https://doi.org/10.1002/14651858.CD003297.pub2

Höttecke, D., & Allchin, D. (2020). Reconceptualizing nature-of-science education in the age of social media. *Science Education*, *104*(4), 641–666. https://doi.org/10.1002/sce.21575

Höttecke, D., & Silva, C. C. (2011). Why Implementing History and Philosophy in School Science Education is a Challenge: An Analysis of Obstacles. *Science & Education*, *20*(3–4), 293–316. https://doi.org/10.1007/s11191-010-9285-4

Howard, S., & Wetter, K. J. (2012). Nanotechnology and Geopolitics: There's Plenty of Room at the Top. In D. Maclurcan & N. Radywyl (Hrsg.), *Nanotechnology and global sustainability* (S. 69–116). CRC Press.

Huber, G. L., & Mandl, H. (1994). *Verbale Daten. Eine Einführung in die Grundlagen und Methoden der Erhebung und Auswertung*. Psychologie Verlags Union.

Hume, D. (2007). *An enquiry concerning human understanding and other writings* (S. Buckle, Hrsg.). Cambridge University press.

Hurlburt, R. T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Consciousness and Cognition*, *22*(4), 1477–1494. https://doi.org/10.1016/j.concog.2013.10.003

Hussy, W., Schreier, M., & Echterhoff, G. (2010). *Forschungsmethoden in Psychologie und Sozialwissenschaften: Für Bachelor*. Springer.

Iacobucci, D., & Churchill, G. A. (2010). *Marketing research: Methodological foundations* (10. Aufl.). South-Western, Cengage Learning.

IBM Deutschland GmbH. *IBM SPSS Statistics*. IBM. Abgerufen 9. November 2024, von https://www.ibm.com/de-de/products/spss-statistics

Ideland, M., Malmberg, C., & Winberg, M. (2011). Culturally Equipped for Socio-Scientific Issues? A comparative study on how teachers and students in mono- and multiethnic schools handle work with complex issues. *International Journal of Science Education*, *33*(13), 1835–1859. https://doi.org/10.1080/09500693.2010.519803

IfD Allensbach (Hrsg.). (2023). *Homöopathie: Nutzung und Wertschätzung in der Bevölkerung*. https://www.ifd-allensbach.de/fileadmin/IfD/sonstige_pdfs/2023_03_15_Pressemitteilung_Homoeopathie.pdf

Institut für Chemiedidaktik (Hrsg.). (2023a). *Modulhandbuch Bachelor of Arts Unterrichtsfacht Chemie—Studienprofil Lehramt an Haupt-, Real-, Sekundar- und Gesamtschulen*. Institut für Chemiedidaktik. https://zfl.uni-koeln.de/sites/zfl/ZfL-Navi/Modulhandbuecher/MNF/Bachelor/MHB_2015_BA_Chemie_LA_HRGe.pdf

Institut für Chemiedidaktik (Hrsg.). (2023b). *Modulhandbuch Bachelor of Arts Unterrichtsfacht Chemie—Studienprofil Lehramt für sonderpädagogische Förderung*. Institut für Chemiedidaktik. https://zfl.uni-koeln.de/sites/zfl/ZfL-Navi/Modulhandbuecher/MNF/Bachelor/MHB_2015_BA_Chemie_LA_SP.pdf

Institut für Chemiedidaktik (Hrsg.). (2023c). *Modulhandbuch Master of Education Unterrichtsfacht Chemie—Studienprofil Lehramt an Haupt-, Real-, Sekundar- und Gesamtschulen*. Institut für Chemiedidaktik. https://zfl.uni-koeln.de/sites/zfl/ZfL-Navi/Modulhandbuecher/MNF/Master/MHB-M_Chemie_HRGe.pdf

Institut für Chemiedidaktik (Hrsg.). (2023d). *Modulhandbuch Master of Education Unterrichtsfacht Chemie—Studienprofil Lehramt für sonderpädagogische Förderung*. Institut für Chemiedidaktik. https://zfl.uni-koeln.de/sites/zfl/ZfL-Navi/Modulhandbuecher/MNF/Master/MHB-M_Chemie_SP.pdf

Janich, P. (1992). *Grenzen der Naturwissenschaft: Erkennen als Handeln* (Originalausgabe). C.H. Beck.

Janssen, J., & Laatz, W. (2017). *Statistische Datenanalyse mit SPSS: Eine anwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-53477-9

Joerger, K. (1976). *Einführung in die Lernpsychologie*. Herder.

Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The Relation Between Culture and Response Styles: Evidence From 19 Countries. *Journal of Cross-Cultural Psychology*, *36*(2), 264–277. https://doi.org/10.1177/0022022104272905

Jülch, V., Hartmann, N., Saad Hussein, N., & Schlegl, T. (2015). Photovoltaik. In M. Wietschel, S. Ullrich, P. Markewitz, F. Schulte, & F. Genoese (Hrsg.), *Energietechnologien der Zukunft: Erzeugung, Speicherung, Effizienz und Netze* (S. 123–138). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-07129-5

Kahneman, D. (1973). *Attention and effort*. Prentice Hall.

Kahneman, D. (2012). *Schnelles Denken, langsames Denken*. Random House GmbH.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases / edited by Daniel Kahneman, Paul Slovic, Amas Tversky*. Cambridge University Press.

Kahneman, D., & Tversky, A. (2008). Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and its Foundations* (S. 114–135). Cambridge University Press.

Kahneman, D., & Tversky, A. (2013). Prospect Theory: An Analysis of Decision Under Risk. In L. C. MacLean & W. T. Ziemba (Hrsg.), *Handbook of the fundamentals of financial decision making* (Bd. 4, S. 99–127). World Scientific Pub. https://doi.org/10.1142/9789814417358_0006

Kanbur, C., Huwer, J., Fetzer, D., Siol, A., & Eilks, I. (2022). Wie man mit Tablets über Seltenerdelemente in Tablets lernen kann. *plusLucis*, *1/2022*, 30–32.

Kant, I. (1787). *Kritik der reinen Vernunft* (Königlich Preußische Akademie der Wissenschaften, Hrsg.; 2. Aufl.). Georg Reimer.

Karle, A. (2015). *Elektromobilität: Grundlagen und Praxis*. Carl Hanser Verlag.

Karn, B., Kuiken, T., & Otto, M. (2009). Nanotechnology and in Situ Remediation: A Review of the Benefits and Potential Risks. *Environmental Health Perspectives*, *117*(12), 1823–1831.

Kaspar, F., & Mächel, H. (2023). Beobachtung von Klima und Klimawandel in Mitteleuropa und Deutschland. In G. P. Brasseur, D. Jacob, & S. Schuck-Zöller (Hrsg.), *Klimawandel in Deutschland: Entwicklung, Folgen, Risiken und Perspektiven* (S. 19–30). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-66696-8

Kelava, A., & Moosbrugger, H. (2020). Testwertverteilung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (3. Aufl., S. 160–170). Springer.

Kessler, G., Veser, A., Schlüter, F.-H., Raskob, W., Landman, C., & Päsler-Sauer, J. (2014). *The Risks of Nuclear Energy Technology: Safety Concepts of Light Water Reactors*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-55116-1

Keuth, H. (2013). Einführung. In H. Keuth (Hrsg.), *Karl Popper: Logik der Forschung* (4., bearb. Aufl., Bd. 12). Akad.-Verl. http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=025787164&sequence=000001&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA

Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching*, *39*(7), 551–578. https://doi.org/10.1002/tea.10036

Klell, M., Eichlseder, H., & Trattner, A. (2018). *Wasserstoff in der Fahrzeugtechnik: Erzeugung, Speicherung, Anwendung* (4. Aufl.). Springer Vieweg.

Klieme, E., Hermann, A., Werner, B., Peter, D., Hans, G., Manfred, P., Kristina, R., Kurt, R., Jürgen, R., Heinz-Elmar, T., & J., V., Helmut. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. https://doi.org/10.25656/01:20901

Knight, A. T., Cowling, R. M., Rouget, M., Balmford, A., Lombard, A. T., & Campbell, B. M. (2008). Knowing but not doing: Selecting priority conservation areas and the research-implementation gap. *Conservation biology : the journal of the Society for Conservation Biology*, *22*(3), 610–617. https://doi.org/10.1111/j.1523-1739.2008.00914.x

Knoblich, G., & Ollinger, M. (2006). Die Methode des lauten Denkens. In J. Funke & P. A. Frensch (Hrsg.), *Handbuch der Allgemeinen Psychologie* (S. 691–696). Kognition Hogrefe.

Knutti, R. (2019). Closing the Knowledge-Action Gap in Climate Change. *One Earth*, *1*(1), 21–23. https://doi.org/10.1016/j.oneear.2019.09.001

Kogut, L. S. (1996). Critical Thinking in General Chemistry. *Journal of Chemical Education*, *73*(3), 218–221.

Köhler, D., Hetzel, M., Klingner, M., Koch, T., Ewig, S., Becher, G., Lindemann, H., Voshaar, T., & Costabel, U. (2019). Die Rolle der Luftschadstoffe für die Gesundheit: Eine Replik auf die Expertise der Internationalen Gesellschaft für Umweltepidemiologie (ISEE) und der European Respiratory Society (ERS). *Pneumologie*, *73*(05), 274–287. https://doi.org/10.1055/a-0873-3574

Kolstø, S. D. (2001). „To trust or not to trust,…'-pupils" ways of judging information encountered in a socio-scientific issue. *International Journal of Science Education*, *23*(9), 877–901. https://doi.org/10.1080/09500690010016102

Kolstø, S. D., Bungum, B., Arnesen, E., Isnes, A., Kristensen, T., Mathiassen, K., Mestad, I., Quale, A., Tonning, A. S. V., & Ulvik, M. (2006). Science students' critical examination of scientific information related to socioscientific issues. *Science Education*, *90*(4), 632–655. https://doi.org/10.1002/sce.20133

Koolwijk, J. von. (1974). Die Befragungsmethode. In J. v. von Koolwijk & M. Wieken-Mayser (Hrsg.), *Techniken der empirischen Sozialforschung* (Bd. 4, S. 9–23). Oldenbourg.

Krause, L.-K., & Gagné, J. (2022, März 3). Corona und Gesellschaft: Wie finden wir aus der Spaltung heraus? | ZEIT ONLINE. *Die Zeit*. https://www.zeit.de/gesellschaft/2022-03/corona-gesellschaft-spaltung-geimpfte-ungeimpfe-moreincommon

Kretschmer, H., & Stary, J. (1998). *Schulpraktikum: Eine Orientierungshilfe zum Lernen und Lehren*. Cornelsen Scriptor.

Kronthaler, F. (2021). *Statistik angewandt mit Excel: Datenanalyse ist (k)eine Kunst*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-62302-2

Krosnick, J. A., Lavrakas, P. J., & Kim, N. (2014). Survey Research. In H. T. Reis & C. M. Judd (Hrsg.), *Handbook of research methods in social and personality psychology* (Bd. 2, S. 404–442). Cambridge University Press.

Krüger, D., Parchmann, I., & Schecker, H. (Hrsg.). (2014). *Methoden in der naturwissenschaftsdidaktischen Forschung*. Springer Berlin Heidelberg.

Kruskal, W. H., & Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Kryjevskaia, M., Heron, P. R. L., & Heckler, A. F. (2021). Intuitive or rational? Students and experts need to be both. *Physics Today*, *74*(8), 28–34. https://doi.org/10.1063/PT.3.4813

Kuckartz, U. (2016). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (3., überarbeitete Aufl.). Beltz.

Kühberger, A. (2016). Framing. In R. F. Pohl (Hrsg.), *Cognitive Illusions: Intriguing Phenomena in Judgement, Thinking and Memory* (2nd ed., S. 79–98). Taylor and Francis. http://gbv.eblib.com/patron/FullRecord.aspx?p=4605464

Kuhn, D. (1999). A Developmental Model of Critical Thinking. *Educational Researcher*, *28*(2), 16. https://doi.org/10.2307/1177186

Kuhn, T. S. (1996). *The structure of scientific revolutions* (3. Aufl.). University of Chicago Press.

Kultusministerkonferenz. (2005). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss—Beschluss vom16.12.2004* (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der & Bundesrepublik Deutschland, Hrsg.). Luchterhand.

Lakatos, I. (2015). Falsification and the Methodology of Scientific Research Programmes'. In I. Lakatos & A. Musgrave (Hrsg.), *Criticism and the growth of knowledge* (Bd. 4, S. 91-196). Cambridge Univ. Press. https://doi.org/10.1017/CBO9781139171434

Lamnek, S., & Krell, C. (2016). *Qualitative Sozialforschung: Mit Online-Material* (6., überarbeitete Aufl.). Beltz.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

Laskar, N., & Kumar, U. (2019). Plastics and microplastics: A threat to environment. *Environmental Technology & Innovation*, *14*, 100352. https://doi.org/10.1016/j.eti.2019.100352

Lawatscheck, C., Hadi, A. N., & Kuse, K.-M. (2022). Das (umwelt)chemische Laborexperiment Neodym – Ein digitales Lernspiel. *plusLucis*, *1/2022*, 27–29.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, *29*(4), 331–359. https://doi.org/10.1002/tea.3660290404

Lederman, N. G. (2018). Nature of Scientific Knowledge and Scientific Inquiry in Biology Teaching. In K. Kampourakis (Hrsg.), *Teaching biology in schools: Global research, issues, and trends* (S. 216–235). Routledge, Taylor & Francis Group.

Lederman, N. G., & Lederman, J. S. (2014). Is Nature of Science Going, Going, Going, Gone? *Journal of Science Teacher Education*, *25*(3), 235–238. https://doi.org/10.1007/s10972-014-9386-z

Lederman, N. G., & Lederman, J. S. (2019). Teaching and learning nature of scientific knowledge: Is it Déjà vu all over again? *Disciplinary and Interdisciplinary Science Education Research*, *1*(1), 6. https://doi.org/10.1186/s43031-019-0002-0

Lee, H., & Yang, J. (2019). Science Teachers Taking their First Steps toward Teaching Socioscientific Issues through Collaborative Action Research. *Research in Science Education*, *49*(1), 51–71. https://doi.org/10.1007/s11165-017-9614-6

Lersch, R. (2007). *Unterricht und Kompetenzerwerb. In 30 Schritten von der Theorie zur Praxis kompetenzfördernden Unterrichts*. https://doi.org/10.25656/01:27309

Lersch, R. (2022). *Wie unterrichtet man Kompetenzen? Didaktik und Praxis kompetenzfördernden Unterrichts*. Hessisches Kultusministerium–Institut für Qualitätsentwicklung.

Lersch, R., & Schreder, G. (2013). *Grundlagen kompetenzorientierten Unterrichtens: Von den Bildungsstandards zum Schulcurriculum*. Verlag Barbara Budrich. https://elibrary.utb.de/doi/book/10.3224/9783847403289

Li, T., & Gao, F. (2024). *Nitrogen fertilizer and wheat: Achieving agricultural production and sustainable development*. Resources Economics Research Board. https://doi.org/10.50908/grb.3.0_28

Liang, L. L., Chen, S., Chen, X., Kaya, O. N., Adams, A. D., Macklin, M., & Ebenezer, J. (2009). Preservice Teachers´Views about Nature of Scientific Knowledge Development: An International Collaborative Study. *International Journal of Science and Mathematics Education*, *7*(5), 987–1012. https://doi.org/10.1007/s10763-008-9140-0

Limesurvey GmbH. (o. J.). *LimeSurvey* (Version 5.6.55+240220) [Web]. LimeSurvey GmbH. Abgerufen 25. Februar 2024, von https://www.limesurvey.org/de

Lipman, M. (1987). Critical thinking: What can it be? *Analytic Teaching*, *8*(1).

Lohmann, D., & Podbregar, N. (2012). *Im Fokus: Bodenschätze: auf der Suche nach Rohstoffen*. Springer.

López-Fernández, M. D. M., González-García, F., & Franco-Mariscal, A. J. (2022). How Can Socio-scientific Issues Help Develop Critical Thinking in Chemistry Education? A Reflection on the Problem of Plastics. *Journal of Chemical Education*, *99*(10), 3435–3442. https://doi.org/10.1021/acs.jchemed.2c00223

Loytved, C., & Schäfers, R. (2023). *Qualitätsbericht 2022, Außerklinische Geburtshilfe in Deutschland* (Gesellschaft für Qualität in der außerklinischen Geburtshilfe e.V., Hrsg.; 1. Aufl.). Verlag Wissenschaftliche Scripten.

Lyman, P., Varian, H. R., Charles, P., Good, N., Jordan, L., & Pal, J. (2003). *How much information? 2003*. University of California at Berkeley, School of Information Science. https://groups.ischool.berkeley.edu/archive/how-much-info-2003/execsum.htm#summary

Maddux, J. E. (1995). Self-Efficacy Theory, An Introduction. In J. E. Maddux (Hrsg.), *Self-Efficacy, Adaptation, and Adjustment: Theory, Research, and Application* (S. 3–33). Springer. https://doi.org/10.1007/978-1-4419-6868-5

Mahner, M. (2013). Science and Pseudoscience - How to Demarcate after the (Alleged) Demise of the Demarcation Problem. In M. Pigliucci & M. Boudry (Hrsg.), *Philosophy of pseudoscience: Reconsidering the demarcation problem* (S. 29–44). The University of Chicago Press.

Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, *18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Mann, R. M., Hyne, R. V., Choung, C. B., & Wilson, Scott. P. (2009). Amphibians and agricultural chemicals: Review of the risks in a complex environment. *Environmental Pollution*, *157*(11), 2903–2927. https://doi.org/10.1016/j.envpol.2009.05.015

Markewitz, P., Bongartz, R., & Biß, K. (2015). Kohlekraftwerke. In M. Wietschel, S. Ullrich, P. Markewitz, F. Schulte, & F. Genoese (Hrsg.), *Energietechnologien der Zukunft: Erzeugung, Speicherung, Effizienz und Netze* (S. 33–56). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-07129-5

Marniok, K. (2018). *Zum Wesen von Theorien und Gesetzen in der Chemie: Begriffsanalyse und Förderung der Vorstellungen von Lehramtsstudierenden*. Logos Verlag Berlin.

Marschall, L., & Holdinghausen, H. (2018). *Seltene Erden: Umkämpfte Rohstoffe des Hightech-Zeitalters*. Oekom Verlag.

Martens, H., & Goldmann, D. (2016). *Recyclingtechnik: Fachbuch für Lehre und Praxis*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-02786-5

Mathie, R. T., Ramparsad, N., Legg, L. A., Clausen, J., Moss, S., Davidson, J. R. T., Messow, C.-M., & McConnachie, A. (2017). Randomised, double-blind, placebo-controlled trials of non-individualised homeopathic treatment: Systematic review and meta-analysis. *Systematic Reviews*, *6*(1), 63. https://doi.org/10.1186/s13643-017-0445-3

Math.-Nat. Fakultät der Universität zu Köln (Hrsg.). (2018). *Modulhandbuch Bachelor of Arts Unterrichtsfacht Chemie—Studienprofil Lehramt an Gymnasien und Gesamtschulen & Studienprofil Lehramt an Berufkollegs*. Math.-Nat. Fakultät der Universität zu Köln. https://zfl.uni-koeln.de/sites/zfl/ZfL-Navi/Modulhandbuecher/MNF/Bachelor/MHB_2018_BA_Chemie_LA_GyGe-BK.pdf

Math.-Nat. Fakultät der Universität zu Köln (Hrsg.). (2021). *Modulhandbuch Master of Education Unterrichtsfacht Chemie—Studienprofil Lehramt an Gymnasien und Gesamtschulen*. Math.-Nat. Fakultät der Universität zu Köln. https://zfl.uni-koeln.de/sites/zfl/ZfL-Navi/Modulhandbuecher/MNF/Master/MHB-M_Chemie_GyGe.pdf

Matt, E. (2019). Darstellung qualitativer Forschung. In U. Flick, E. Kardorff, & I. Steinke (Hrsg.), *Qualitative Forschung: Ein Handbuch* (13. Aufl., Originalausgabe, Bd. 55628, S. 578–586). Rowohlts Enzyklopädie im Rowohlt Taschenbuch Verlag.

Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (11. Neuausgabe). Beltz. http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1143991

Mayring, P. (2019). Qualitative Inhaltsanalyse. In U. Flick, E. Kardorff, & I. Steinke (Hrsg.), *Qualitative Forschung: Ein Handbuch* (13. Aufl., Originalausgabe, Bd. 55628, S. 468–474). rowohlts enzyklopädie im Rowohlt Taschenbuch Verlag.

Mayring, P. (2020). Qualitative Inhaltsanalyse. In G. Mey & K. Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie: Band 2: Designs und Verfahren* (2., überarb. Aufl. 2020). Springer Fachmedien Wiesbaden GmbH.

McComas, W. F. (1998a). *The Nature of Science in Science Education: Rationales and Strategies* (Bd. 5). Springer Netherlands.

McComas, W. F. (1998b). The Principal Elements of the Nature of Science: Dispelling the Myths. In W. F. McComas (Hrsg.), *The nature of science in science education: Rationales and strategies* (S. 53–70). Kluwer Academic Publishers.

McComas, W. F. (2020). Principal Elements of Nature of Science: Informing Science Teaching while Dispelling the Myths. In W. F. McComas (Hrsg.), *Nature of science in science instruction: Rationales and strategies* (S. 35–65). Springer.

McComas, W. F., Clough, M. P., & Nouri, N. (2020). Nature of Science and Classroom Practice: A Review of the Literature with Implications for Effective NOS Instruction. In W. F. McComas (Hrsg.), *Nature of science in science instruction: Rationales and strategies* (S. 67–114). Springer.

McComas, W. F., & Olson, J. K. (1998). The Nature of Science in International Science Education Standards Documents. In W. F. McComas (Hrsg.), *The nature of science in science education: Rationales and strategies* (S. 41–52). Kluwer Academic Publishers.

McNeil, S. E. (2011). Unique Benefits of Nanotechnology to Drug Delivery and Diagnostics. In S. E. McNeil (Hrsg.), *Characterization of Nanoparticles Intended for Drug Delivery* (Bd. 697, S. 1–8). Humana Press. https://doi.org/10.1007/978-1-60327-198-1

Mercier, H. (2016). Confirmation bias – myside bias. In R. F. Pohl (Hrsg.), *Cognitive Illusions: Intriguing Phenomena in Judgement, Thinking and Memory* (2. Aufl. S. 99–114). Taylor and Francis. http://gbv.eblib.com/patron/FullRecord.aspx?p=4605464

Mesci, G., & Schwartz, R. S. (2017). Changing Preservice Science Teachers' Views of Nature of Science: Why Some Conceptions May be More Easily Altered than Others. *Research in Science Education*, *47*(2), 329–351. https://doi.org/10.1007/s11165-015-9503-9

Mesquita, B., Frijda, N., & Scherer, K. (1997). Culture and Emotion. In J. W. Berry, P. R. Dasen, & T. S. Saraswati (Hrsg.), *Basic processes and human development* (2. Aufl., S. 255–298). Allyn and Bacon.

Meßmer, A.-K., Sängerlaub, A., & Schulz, L. (2021). *„Quelle: Internet"? Digitale Nachrichten- und Informationskompetenzen der deutschen Bevölkerung im Test* (Stiftung Neue Verantwortung e.V., Hrsg.). Hochschule der Medien.

Mestre, J. P., & Cocking, R. R. (2002). Applying the Science of Learning to the Education of Prospective Science Teachers. In R. W. Bybee (Hrsg.), *Learning science and the science of learning*. NSTA Press.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*(1), 3–19. https://doi.org/10.1037/0033-295X.106.1.3

Mey, G., & Mruck, K. (Hrsg.). (2020). *Handbuch Qualitative Forschung in der Psychologie: Band 2: Designs und Verfahren* (2., erw. u. überarb. Auflage 2020). Springer Fachmedien Wiesbaden GmbH.

Meyer, H. L. (1977). *Trainingsprogramm zur Lernzielanalyse* (6., erw. Aufl). Athenäum.

Microsoft Corporation. *Microsoft Excel*. Microsoft. Abgerufen 09. November 2024, von https://www.microsoft.com/de-de/microsoft-365/excel?market=de

Mieger, D., & Prechtl, M. (2017). Recycling von Neodym(III)-sulfat aus Festplattenmagneten: Seltenerdelemente im Schülerforschungsprojekt und im Berufsschulunterricht. *Naturwissenschaften im Unterricht. Chemie*, *28*(161), 39–43.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (2019). *Kernlehrplan für die Sekundarstufe I Gymnasium in Nordrhein-Westfalen: Bd. Nr. 3415* (1. Auflage).

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (2022). *Kernlehrplan für die Sekundarstufe II Gymnasium, Gesamtschule in Nordrhein-Westfalen: Bd. Nr. 4723* (1. Aufl.).

Minott, D., Ferguson, T., & Minott, G. (2019). Critical Thinking and Sustainable Development. In W. Leal Filho (Hrsg.), *Encyclopedia of Sustainability in Higher Education* (S. 1–6). Springer International Publishing. https://doi.org/10.1007/978-3-319-63951-2_529-1

Mohr, H. (1989). Verfügungswissen und Orientierungswissen: Die Verantwortung des Wissenschaftlers. *Mathematisch-Naturwissenschaftlicher Unterricht*, *42*, 127–132.

Mohrdiek, C., Venturi, M., & Breitrück, K. (2017). Mobile Anwendungen. In J. Töpler, J. Lehmann, & E. U. von Weizsäcker (Hrsg.), *Wasserstoff und Brennstoffzelle: Technologien und Marktperspektiven* (2., aktualisierte und erweiterte Auflage, S. 59–114). Springer Vieweg. https://doi.org/10.1007/978-3-662-53360-4

Monat, J. P., & Gannon, T. F. (2015). What is Systems Thinking? A Review of Selected Literature Plus Recommendations. *American Journal of Systems Science*, *4*(1), 11–26. https://doi.org/10.5923/j.ajss.20150401.02

Moors, G., Kieruj, N. D., & Vermunt, J. K. (2014). The Effect of Labeling and Numbering of Response Scales on the Likelihood of Response Bias. *Sociological Methodology*, *44*(1), 369–399. https://doi.org/10.1177/0081175013516114

Moosbrugger, H., & Brandt, H. (2020a). Antwortformate und Itemtypen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (3., vollständig neu bearbeitete, erweiterte und aktualisierte Auflage, S. 93–118). Springer.

Moosbrugger, H., & Brandt, H. (2020b). Itemkonstruktion und Antwortverhalten. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (3., vollständig neu bearbeitete, erweiterte und aktualisierte Auflage, S. 67–90). Springer.

Moosbrugger, H., & Brandt, H. (2020c). Planungsaspekte und Konstruktionsphasen von Tests und Fragebogen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (3., vollständig neu bearbeitete, erweiterte und aktualisierte Auflage, S. 39–66). Springer.

Moosbrugger, H., & Kelava, A. (2020a). Deskriptivstatistische Itemanalyse und Testwertbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (3., vollständig neu bearbeitete, erweiterte und aktualisierte Auflage, S. 143–158). Springer.

Moosbrugger, H., & Kelava, A. (2020b). Qualitätsanforderungen an Tests und Fragebogen („Gütekriterien"). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (3., vollständig neu bearbeitete, erweiterte und aktualisierte Auflage, S. 13–38). Springer.

Mortimer, C. E., & Müller, U. (Hrsg.). (2010). *Chemie* (10. Aufl.). Georg Thieme Verlag. https://doi.org/10.1055/b-002-54081

Müller, L. (2021). *Alltagsentscheidungen für den Chemieunterricht erkennen und Entscheidungsprozesse explorativ begleiten* (Bd. 318). Logos Verlag.

Müller, S. (2021). *Die Vorläufigkeit und soziokulturelle Eingebundenheit naturwissenschaftlicher Erkenntnisse* (Bd. 317). Logos Verlag.

Müller, S., & Reiners, C. S. (2021). Is Chemical Knowledge Tentative or Durable or Both? – Analysis of Preservice Science Teachers' Conceptions about the Tentativeness of Chemical Knowledge. *CHEMKON*, *28*(5), 187–195. https://doi.org/10.1002/ckon.201900080

Nahrstedt, H. (2017). *Excel in Perfektion*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-19441-3

Napper, I. E., Bakir, A., Rowland, S. J., & Thompson, R. C. (2015). Characterisation, quantity and sorptive properties of microplastics extracted from cosmetics. *Marine Pollution Bulletin*, *99*(1–2), 178–185. https://doi.org/10.1016/j.marpolbul.2015.07.029

Nathaniel, S. P., Alam, Md. S., Murshed, M., Mahmood, H., & Ahmad, P. (2021). The roles of nuclear energy, renewable energy, and economic growth in the abatement of carbon dioxide emissions in the G7 countries. *Environmental Science and Pollution Research*, *28*(35), 47957–47972. https://doi.org/10.1007/s11356-021-13728-6

National Academies of Sciences, Engineering, and Medicine (Hrsg.). (2016). *Science literacy: Concepts, contexts, and consequences*. National Academies Press.

National Research Council. (1996). *National Science Education Standards*. National Academy Press.

National Science Board. (2018). *Science & Engineering Indicators 2018*. https://www.nsf.gov/statistics/2018/nsb20181/report/sections/science-and-technology-public-attitudes-and-understanding/notes

Neumann, I., & Kremer, K. (2013). Nature of Science und epistemologische Überzeugungen: Ähnlichkeiten und Unterschiede. *Zeitschrift für Didaktik der Naturwissenschaften*, *19*, 209–232.

Niebert, K., & Gropengießer, H. (2014). Leitfadengestützte Interviews. In D. Krüger, I. Parchmann, & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 121–132). Springer Berlin Heidelberg.

Nobach, H. (2023). *Practical Realization of Bessel's Correction for a Bias-Free Estimation of the Auto-Covariance and the Cross-Covariance Functions* (arXiv:2303.11047). arXiv. http://arxiv.org/abs/2303.11047

Nordmann, J., Welfens, M. J., Fischer, D., Nemnich, C., Bookhagen, B., Bienge, K., & Niebert, K. (2015). *Die Rohstoff-Expedition: Entdecke, was in (d)einem Handy steckt*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-44083-4

Olson, D. R., & Astington, J. W. (1993). Thinking About Thinking: Learning How to Take Statements and Hold Beliefs. *Educational Psychologist*, *28*(1), 7–23. https://doi.org/10.1207/s15326985ep2801_2

Organisation for Economic Co-operation and Development (Hrsg.). (1999). *Measuring student knowledge and skills: A new framework for assessment*. OECD.

Organisation for Economic Co-operation and Development (Hrsg.). (2010). *The Definition and Selection of Key Competencies: Executive Summary*. OECD.

Osborne, J. (2014). Teaching critical thinking? New directions in science education. *School Science Review*, *352*, 53–62.

Osborne, J. (2023). Science, Scientific Literacy and Science Education. In N. G. Lederman, D. L. Zeidler, & J. S. Lederman (Hrsg.), *Handbook of research on science education. Vol. 3* (Bd. 3, S. 755–816). Routledge.

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What „ideas-about-science" should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, *40*(7), 692–720. https://doi.org/10.1002/tea.10105

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, *11*(6), 988–1010. https://doi.org/10.3758/BF03196730

Pappenheim, E. (1892). Johann Amos Comenius - I. Teil. In Dr. G. Fröhlich (Hrsg.), *Die Klassiker der Pädagogik* (2. Aufl., Bd. 15). Schulbuchhandlung von F. G. L Gressler.

Paul, R., & Elder, L. (2014). *Critical thinking: Tools for taking charge of your learning and your life* (3. Aufl., Pearson new international edition). Pearson Education Limited.

Paul, R., & Elder, L. (2019a). *The thinker's guide: For conscientious citizens on how to detect media bias & propaganda in national and world news* (4. Aufl.). The Foundation for Critical Thinking.

Paul, R., & Elder, L. (2019b). *The Thinker's Guide to Fallacies: The Art of Mental Trickery and Manipulation*. Rowman & Littlefield Publishers.

Perkins, D. N. (1987). Conversation with David N. Perkins. In M. Heiman & J. Slomianko, *Thinking skills instruction: Concepts and techniques* (S. 49–57). National Education Association.

Peters, R. (1966). *Ethics in Education*. Allen and Unwin.

Petri, J. (2014). Fallstudien zur Analyse von Lernpfaden. In D. Krüger, I. Parchmann, & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 95–106). Springer Berlin Heidelberg.

Pfeiffer, E.-M., Eschenbach, A., Munch, J. C., & Vereecken, H. (2023). Böden und ihre Funktion im Klimawandel. In G. P. Brasseur, D. Jacob, & S. Schuck-Zöller (Hrsg.), *Klimawandel in Deutschland: Entwicklung, Folgen, Risiken und Perspektiven* (S. 262–272). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-66696-8

Pithers, R. T., & Soden, R. (2000). Critical thinking in education: A review. *Educational Research*, *42*(3), 237–249. https://doi.org/10.1080/001318800440579

Pohl, R. F. (Hrsg.). (2004). *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory* (1. Aufl.). Psychology Press. http://www.loc.gov/catdir/enhancements/fy0653/2004009834-d.html

Popper, K. R. (1935). *Logik der Forschung*. Springer Vienna. https://doi.org/10.1007/978-3-7091-4177-9

Popper, K. R. (1976). *Logik der Forschung* (6. Aufl.). Mohr.

Porst, R. (2014). *Fragebogen: Ein Arbeitsbuch*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-02118-4

Prechtl, M. (2017). Cer sorgt für den zündende Funken im Feuerzeug: Schulversuche mit einem Seltenerdelement. *Naturwissenschaften im Unterricht. Chemie*, *28*(161), 32–34.

Prechtl, M. (2022). Seltenerdelemente. *plusLucis*, *1/2022*, 4–10.

Prechtl, M., & Legscha, Y. L. (2022). Storytelling zu Seltenerdelementen. *plusLucis*, *1/2022*, 11–15.

Prenzel, M., Rost, J., Senkbeil, M., Häußler, P., & Klopp, A. (2001). Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann, & M. Weiß (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 191–248). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-83412-6_6

Primas, H. (1985). Kann Chemie auf Physik reduziert werden?? Erster Teil: Das Molekulare Programm. *Chemie in unserer Zeit*, *19*(4), 109–119. https://doi.org/10.1002/ciuz.19850190402

Pronin, E., Lin, D. Y., & Ross, L. (2002). The Bias Blind Spot: Perceptions of Bias in Self Versus Others. *Personality and Social Psychology Bulletin*, *28*(3), 369–381.

Przyborski, A., & Wohlrab-Sahr, M. (2021). *Qualitative Sozialforschung: Ein Arbeitsbuch* (5., überarbeitete und erweiterte Auflage). De Gruyter Oldenbourg.

Rafolt, S., Kapelari, S., & Kremer, K. (2019). Kritisches Denken im naturwissenschaftlichen Unterricht – Synergiemodell, Problemlage und Desiderata. *Zeitschrift für Didaktik der Naturwissenschaften*, *25*(1), 63–75. https://doi.org/10.1007/s40573-019-00092-9

Räsänen, A., Sarkki, S., Haanpää, O., Isolahti, M., Kekkonen, H., Kikuchi, K., Koukkari, V., Kärkkäinen, K., Miettinen, J., Mäntymaa, E., Nieminen, M., Rahkila, R., Ruohonen, A., Sarkkola, S., Välimäki, M., Yliperttula, K., & Heikkinen, H. I. (2024). Bridging the knowledge-action gap: A framework for co-producing actionable knowledge. *Environmental Science & Policy*, *162*, 103929. https://doi.org/10.1016/j.envsci.2024.103929

Rasch, B., Friese, M., Hofmann, W., & Naumann, E. (2014a). *Quantitative Methoden 1: Einführung in die Statistik für Psychologen und Sozialwissenschaftler*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-43524-3

Rasch, B., Friese, M., Hofmann, W., & Naumann, E. (2014b). *Quantitative Methoden 2*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-43548-9

Rat der Europäischen Union. (2008). Amtsblatt der Europäischen Union: Richtlinie 2008/50/EG des europäischen Parlaments und des Rates vom 21. Mai 2008 über Luftqualität und saubere Luft für Europa. *Amtsblatt der Europäischen Union*, *50*(L 152). https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32008L0050

Rat der Europäischen Union. (2017). Amtsblatt der Europäischen Union: Mitteilungen und Bekanntmachungen. *Amtsblatt der Europäischen Union*, *60*(C 189). https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=OJ:C:2017:189:FULL&from=CS

Rat der Europäischen Union. (2018). Empfehlungen des Rates vom 22. Mai 2018 zu Schlüsselkompetenzen für lebenslanges Lernen: (Text von Bedeutung für den EWR). *Amtsblatt der Europäischen Union*, *C 189*, 1–13.

Reber, A. S. (1989). Implicit leraning and tactic knowledge. *Journal of Experimental Psychology*, *118*(3), 219.

Rechid, D., Breil, M., Jacob, D., Kottmeier, C., Petersen, J., Pfeifer, S., & Teichmann, C. (2023). Regionale Klimamodellierung. In G. P. Brasseur, D. Jacob, & S. Schuck-Zöller (Hrsg.), *Klimawandel in Deutschland: Entwicklung, Folgen, Risiken und Perspektiven* (S. 31–46). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-66696-8

Reganold, J. P., & Wachter, J. M. (2016). Organic agriculture in the twenty-first century. *Nature Plants*, *2*(2), 15221. https://doi.org/10.1038/nplants.2015.221

Reich, K. (Hrsg.). (2017). Methodenpool. *http://methodenpool.uni-koeln.de*.

Reichertz, J. (with Springer Fachmedien Wiesbaden). (2016). *Qualitative und interpretative Sozialforschung: Eine Einladung*. Springer VS.

Reinders, H. (2022a). Fragebogen. In H. Reinders, D. Bergs-Winkels, A. Prochnow, & I. Post (Hrsg.), *Empirische Bildungsforschung: Eine elementare Einführung* (S. 161–173). Springer VS. https://doi.org/10.1007/978-3-658-27277-7

Reinders, H. (2022b). Interview. In H. Reinders, D. Bergs-Winkels, A. Prochnow, & I. Post (Hrsg.), *Empirische Bildungsforschung: Eine elementare Einführung* (S. 211–222). Springer VS. https://doi.org/10.1007/978-3-658-27277-7

Reinders, H., & Bergs-Winkels, D. (2022). Methodische Fehler im Forschungsprozess. In H. Reinders, D. Bergs-Winkels, A. Prochnow, & I. Post (Hrsg.), *Empirische Bildungsforschung: Eine elementare Einführung* (S. 281–316). Springer VS. https://doi.org/10.1007/978-3-658-27277-7

Reinders, H., Bergs-Winkels, D., Prochnow, A., & Post, I. (Hrsg.). (2022). *Empirische Bildungsforschung: Eine elementare Einführung*. Springer VS. https://doi.org/10.1007/978-3-658-27277-7

Reinders, H., & Post, I. (2022). Quantitative Auswertungsverfahren. In H. Reinders, D. Bergs-Winkels, A. Prochnow, & I. Post (Hrsg.), *Empirische Bildungsforschung: Eine elementare Einführung* (S. 269–280). Springer VS. https://doi.org/10.1007/978-3-658-27277-7

Reiners, C. S. (1996). Von der Stoffsystematik zu systemischen Aspekten der Naturwissenschaft Chemie: Zur Konzeption eines Praktikums für Lehramtskandidatinnen und -kandidaten, das zu komplexem Denken und entsprechend differenziertem Urteil hinführen soll. *Zeitschrift für Didaktik der Naturwissenschaften*, *2*(1), 11–20.

Reiners, C. S. (2003). Die Scandiumgruppe (3. Gruppe, III. Nebengruppe). In W. Glöckler, W. Jansen, & R. G. Weissenhorn (Hrsg.), *Handbuch der experimentellen Chemie Sek. II. Bd.5: Chemie der gebrauchsmetalle* (Bd. 5, S. 1–27). Aulis.

Reiners, C. S. (2017). Kritische Rohstoffe! - Kritische Unterrichtsstoffe?: Eine fachdidaktische Reflexion am Beispiel der Seltenen Erden. *Naturwissenschaften im Unterricht. Chemie*, *28*(161), 8–12.

Reiners, C. S. (2022). *Chemie vermitteln: Fachdidaktische Grundlagen und Implikationen* (2. erweiterte Auflage). Springer Spektrum. https://doi.org/10.1007/978-3-662-65511-5

Riese, J., & Reinhold, P. (2014). Entwicklung eines Leistungstests für fachdidaktisches Wissen. In D. Krüger, I. Parchmann, & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 257–267). Springer Berlin Heidelberg.

Roberts, R. M. (1989). *Serendipity: Accidental discoveries in science*. Wiley.

Roche, D. G., O'Dea, R. E., Kerr, K. A., Rytwinski, T., Schuster, R., Nguyen, V. M., Young, N., Bennett, J. R., & Cooke, S. J. (2022). Closing the knowledge-action gap in conservation with open science. *Conservation biology: the journal of the Society for Conservation Biology*, *36*(3), 13835. https://doi.org/10.1111/cobi.13835

Roeser, S. (2011). Nuclear Energy, Risk, and Emotions. *Philosophy & Technology*, *24*(2), 197–201. https://doi.org/10.1007/s13347-011-0027-6

Rosenthal, J. (2009). Induktion und Bestätigung. In A. Bartels & M. Stöckler (Hrsg.), *Wissenschaftstheorie: Ein Studienbuch* (2., durchges. und korrigierte Aufl, S. 109–133). mentis-Verl.

Röth, T., Kampker, A., Reisgen, U., Hameyer, K., De Docker, R. W., Stolze, T., Vetter, A., Hagedorn, J., Sauer, D. U., & Müller, D. (2013). Entwicklung von elektrofahrzeugspezifischen Systemen. In A. Kampker, D. Vallée, & A. Schnettler (Hrsg.), *Elektromobilität: Grundlagen einer Zukunftstechnologie* (S. 235–334). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31986-0

Rott, H. (2008). Believe Revision. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and its Foundations*. Cambridge University Press.

Rottener, C., Jarman, C., Sutton, S., & Sayres, S. (2024). Size Onset of Metallic Behavior in Neutral Aluminum Clusters. *Nanoscale*, *28*(16), 13516–13524. https://doi.org/10.26434/chemrxiv-2024-9dtb7

Röttgers, K. (1975). *Kritik und Praxis: Zur Geschichte des Kritikbegriffs von Kant bis Marx*. de Gruyter.

Russel, B. (1922, März 24). *Free Thought and Official Propaganda* [Lecture]. Conway Memorial Lecture, South Place Institute, London, GB.

Rychen, D. S., & Salganik, L. H. (Hrsg.). (2001). *Defining and selecting key competencies*. Hogrefe & Huber.

Sächsisches Staatsministerium für Kultus. (2022). *Lehrplan Gymnasium: Chemie* (Landesamt für Schule und Bildung, Hrsg.)

Sadler, T. D. (2011). Situating Socio-scientific Issues in Classrooms as a Means of Achieving Goals of Science Education. In T. D. Sadler (Hrsg.), *Socio-scientific Issues in the Classroom* (Bd. 39, S. 1–10). Springer Netherlands. https://doi.org/10.1007/978-94-007-1159-4

Sadler, T. D., Barab, S. A., & Scott, B. (2007). What Do Students Gain by Engaging in Socioscientific Inquiry? *Research in Science Education*, *37*(4), 371–391. https://doi.org/10.1007/s11165-006-9030-9

Sadler, T. D., Klostermann, M. L., & Topcu, M. S. (2011). Learning Science Content and Socioscientific Reasoning Through Classroom Explorations of Global Climate Change. In T. D. Sadler (Hrsg.), *Socio-scientific Issues in the Classroom* (Bd. 39, S. 45–78). Springer Netherlands. https://doi.org/10.1007/978-94-007-1159-4

Salancik, G. R., & Pfeffer, J. (1977). An Examination of Need-Satisfaction Models of Job Attitudes. *Administrative Science Quarterly*, *22*(3), 427. https://doi.org/10.2307/2392182

Sämmer, G. (1999). Psychologische Gesetzmäßigkeiten bei Beurteilungen in der Schule. *Seminar*, *2*.

Sander, H., & Höttecke, D. (2017). Bewertungskompetenz in der Physikdidaktik: Zwischen Rationalität und Intuition. In *Fachlich argumentieren lernen. Didaktische Forschungen zur Argumentation in den Unterrichtsfächern*. Waxmann. https://www.pedocs.de/frontdoor.php?source_opus=14021

Sandmann, A. (2014). Lautes Denken – die Analyse von Denk-, Lern- und Problemlöseprozessen. In D. Krüger, I. Parchmann, & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 179–188). Springer Berlin Heidelberg.

Sauer, D. U., & Ringbeck, F. (2024). Batteriesysteme und deren Steuerung. In A. Kampker & H. H. Heimes (Hrsg.), *Elektromobilität: Grundlagen einer Fortschrittstechnologie* (S. 129–152). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-65812-3

Schaefer, G. (2002). Scientific Literacy, General Education and Implications for Teaching „Subject-Transcendent Subjects" in School. In *Qualitätsentwicklung im naturwissenschaftlichen Unterricht. Fachtagung am 15. Dezember 1999: Bd. Materialien zur Bildungsforschung*. GFPF; DIPF: Frankfurt am Main.

Schmidt, H., Eyring, V., Latif, M., Marotzke, J., Rechid, D., & Sausen, R. (2023). Globale Modellierung des Klimawandels. In G. P. Brasseur, D. Jacob, & S. Schuck-Zöller (Hrsg.), *Klimawandel in Deutschland: Entwicklung, Folgen, Risiken und Perspektiven* (S. 7–18). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-66696-8

Schmidtchen, U., & Wurster, R. (2017). Sicherheit in der Anwendung von Wasserstoff. In J. Töpler, J. Lehmann, & E. U. von Weizsäcker (Hrsg.), *Wasserstoff und Brennstoffzelle: Technologien und Marktperspektiven* (2., aktualisierte und erweiterte Auflage, S. 43–58). Springer Vieweg. https://doi.org/10.1007/978-3-662-53360-4

Schnabel, J. (2020). Atomausstieg 2022, Konflikte beim Rückbau von Kernkraftwerken in Baden-Württemberg und Schleswig-Holstein. In F. Brettschneider (Hrsg.), *Bau- und Infrastrukturprojekte: Dialogorientierte Kommunikation als Erfolgsfaktor* (S. 333–420). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-28235-6

Schneiders, W. (1985). Vernünftiger Zweifel und wahre Eklektik. Zur Entstehung des modernen Kritikbegriffes. *Studia Leibnitiana*, *17*(2), 143–161.

Scholz, O. R., & Schmacke, N. (2022). *Münsteraner Memorandum Anthroposophische Medizin*. https://muensteraner-kreis.de/wp-content/uploads/2022/12/Muensteraner-Memorandum-Anthroposophische-Medizin.pdf

Schöne-Seifert, B. (Hrsg.). (2023). *Münsteraner Memoranden: Eine kritische Auseinandersetzung mit komplementärer und alternativer Medizin* (1. Aufl.). jmb-Verlag.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting Self-Regulation in Science Education: Metacognition as Part of a Broader Perspective on Learning. *Research in Science Education*, *36*(1–2), 111–139. https://doi.org/10.1007/s11165-005-3917-8

Schultz, R., Skoumal, R. J., Brudzinski, M. R., Eaton, D., Baptie, B., & Ellsworth, W. (2020). Hydraulic Fracturing-Induced Seismicity. *Reviews of Geophysics*, *58*(3), 1-43. https://doi.org/10.1029/2019RG000695

Schumacher, A. (2015). *Paving the Way Towards Authentic Chemistry Teaching: A Contribution to Teachers' Professional Development* (Bd. 184). Logos Verlag Berlin.

Seifert, G., Blakeslee, S. B., Calaminus, G., Kandil, F. I., Barth, A., Bernig, T., Classen, C. F., Corbacioglu, S., Föll, J., Gottschling, S., Gruhn, B., Vom Hoff-Heise, C., Lode, H. N., Martin, D., Nathrath, M., Neunhoeffer, F., Pekrun, A., Wulff, B., Zuzak, T., Henze, G. & Längler, A. (2022). Integrative medicine during the intensive phase of chemotherapy in pediatric oncology in Germany: A randomized controlled trial with 5-year follow up. *BMC Cancer*, *22*(1), 652. https://doi.org/10.1186/s12885-022-09703-0

Sekretariat der Kultusministerkonferenz. (2016). *Strategie der Kultusministerkonferenz „Bildung in der digitalen Welt"*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2016/2016_12_08-Bildung-in-der-digitalen-Welt.pdf

Sekretariat der Kultusministerkonferenz. (2019). *Standards für die Lehrerbildung: Bildungswissenschaften: Beschluss der Kultusministerkonferenz vom 16.12.2004 i. D. F. vom 16.05.2019*. https://www.zfsl.nrw.de/mue/seminar_gyge/seminarprogramm/25_standards-bildungswissenschaften-kmk-2019-0516.pdf

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2024a). *Bildungsstandards Naturwissenschaften (2024) Biologie, Chemie, Physik—Sekundarstufe I Beitrag zur Implementation*. Carl Link Verlag.

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2024b). *Weiterentwickelte Bildungsstandards in den Naturwissenschaften für das Fach Chemie (MSA)*. Carl Link Verlag.

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2020). *Bildungsstandards im Fach Chemie für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 18.06.2020* (1. Aufl.). Carl Link Verlag.

Shaked, H., & Schechter, C. (2017). *Systems Thinking for School Leaders*. Springer International Publishing. https://doi.org/10.1007/978-3-319-53571-5

Shamos, M. H. (1995). *The myth of scientific literacy*. Rutgers Univ. Press.

Shamos, M. H. (2002). Durch Prozesse ein Bewußtsein für die Naturwissenschaften entwickeln. In W. Gräber, P. Nentwig, T. Koballa, & R. Evans (Hrsg.), *Scientific Literacy*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-80863-9

Sharon, A. J., & Baram-Tsabari, A. (2020). Can science literacy help individuals identify misinformation in everyday life? *Science Education*, *104*(5), 873–894. https://doi.org/10.1002/sce.21581

Shechtman, D. (2021). Quasi-Crystal, Not Quasi-Scientist. *Frontiers for Young Minds*, *8*, 22. https://doi.org/10.3389/frym.2020.00022

Shen, L., & Worrell, E. (2024). Plastic recycling Li Shen and Ernst Worrell. In C. Meskers, E. Worrell, & M. A. Reuter (Hrsg.), *Handbook of recycling: State-of-the-art for practitioners, analysts, and scientists* (Second edition, S. 497–510). Elsevier.

Siebenpfeiffer, W. (Hrsg.). (2021). *Mobilität der Zukunft: Intermodale Verkehrskonzepte*. Springer Vieweg.

Silver, S. (2001). Trial results warn of dangers in the use of mistletoe extract. *The Lancet Oncology*, *2*(4), 196.

Silvia, P. J. (2021). The self-reflection and insight scale: Applying item response theory to craft an efficient short form. *Current Psychology*, *41*(12), 8635–8645. https://doi.org/10.1007/s12144-020-01299-7

Silvia, P. J., & Phillips, A. G. (2011). Evaluating self-reflection and insight as self-conscious traits. *Personality and Individual Differences*, *50*, 234–237. https://doi.org/10.1016/j.paid.2010.09.035

Simon, H. A. (1992). What is an "Explanation" of Behavior? *Psychological Science*, *3*(3).

Simpson, M., Parsons, M., Greenwood, J., & Wade, K. (2001). Raspberry Leaf in Pregnancy: Its Saftey and Efficacy in Labor. *Journal of Midwifery & Women's Health*, *46*(2), 51–59. https://doi.org/10.1016/S1526-9523(01)00095-2

Sinatra, G. M., Kienhues, D., & Hofer, B. K. (2014). Addressing Challenges to Public Understanding of Science: Epistemic Cognition, Motivated Reasoning, and Conceptual Change. *Educational Psychologist*, *49*(2), 123–138. https://doi.org/10.1080/00461520.2014.916216

Sinclair, M. (2010). Misconceptions About Intuition. *Psychological Inquiry*, *21*(4), 378–386. https://doi.org/10.1080/1047840X.2010.523874

Siqueira, D. S., De Almeida Meystre, J., Hilário, M. Q., Rocha, D. H. D., Menon, G. J., & Da Silva, R. J. (2019). Current perspectives on nuclear energy as a global climate change mitigation option. *Mitigation and Adaptation Strategies for Global Change*, *24*(5), 749–777. https://doi.org/10.1007/s11027-018-9829-5

Sjöström, J., & Eilks, I. (2020). Reconsidering Different Visions of Scientific Literacy and Science Education Based on the Concept of Bildung. In Y. J. Dori, Z. R. Mevarech, & D. R. Baker (Hrsg.), *Cognition, Metacognition, and Culture in STEM Education* (Bd. 24, S. C1–C1). Springer International Publishing. https://doi.org/10.1007/978-3-319-66659-4_16

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The Affect Heuristic. In T. Gilovich, D. Griffin, & D. Kahneman (Hrsg.), *Heuristics and biases: The psychology of intuitive judgement* (S. 397–420). Cambridge University Press.

Smeriglio, A., Tomaino, A., & Trombetta, D. (2014). Herbal Products in Pregnancy: Experimental Studies and Clinical Reports: Herbal Products in Pregnancy. *Phytotherapy Research*, *28*(8), 1107–1116. https://doi.org/10.1002/ptr.5106

Smidt, W. (2017). Frühe Bildung, Erziehung und Betreuung im Wandel der Zeit – Einige Anmerkungen zur Entwicklung der Elementarpädagogik an der Leopold-Franzens-Universität Innsbruck. In L. Burkhardt & J. C. Aigner (Hrsg.), *Kritische Lektionen* (1. Auflage). Innsbruck University Press.

Stanovich, K. E. (2008). Individual Differences in Reasoning and the Algorithmic/Intentional Level Distinction in Cognitive Science. In J. E. Adler & L. J. Rips, *Reasoning: Studies of Human Inference and its Foundations* (S. 414–436). Cambridge University Press.

Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, *16*(1), 27–44. https://doi.org/10.1016/0030-5073(76)90005-2

Stein, D., & Grant, A. M. (2014). Disentangling the Relationships Among Self-Reflection, Insight, and Subjective Well-Being: The Role of Dysfunctional Attitudes and Core Self-Evaluations. *The Journal of Psychology*, *148*(5), 505–522. https://doi.org/10.1080/00223980.2013.810128

Steinke, I. (2019). Gütekriterien qualitativer Forschung. In U. Flick, E. Kardorff, & I. Steinke (Hrsg.), *Qualitative Forschung: Ein Handbuch* (13. Aufl., Originalausgabe, Bd. 55628, S. 319–331). Rowohlts Enzyklopädie im Rowohlt Taschenbuch Verlag.

Ströker, E. (1968). *Falsifizierbarkeit als Kennzeichen naturwissenschaftlicher Theorien (zu K. R. Poppers „Logik der Forschung"):* Walter de Gruyter Verlag.

Talens Peiró, L., & Villalba Méndez, G. (2013). Material and Energy Requirement for Rare Earth Production. *JOM*, *65*(10), 1327–1340. https://doi.org/10.1007/s11837-013-0719-8

Taran, M., Safaei, M., Karimi, N., & Almasi, A. (2020). Benefits and Application of Nanotechnology in Environmental Science: An Overview. *Biointerface Research in Applied Chemistry*, *11*(1), 7860–7870. https://doi.org/10.33263/BRIAC111.78607870

Taylor, D., Rogers, A. L., & Veal, W. R. (2009). Using Self-Reflection to Increase Science Process Skills in the General Chemistry Laboratory. *Journal of Chemical Education*, *86*(3), 393. https://doi.org/10.1021/ed086p393

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences* (Nachdr.). Sage Publishing.

The Independent. (2018, Mai 12). *Trump cancels NASA's greenhouse gas monitoring system after scrapping clean air regulation*. The Independent.

Thomas, G., & Durant, J. (1987). *Why should we promote the public understanding of science. 1*, 1–14. University of Oxford, Department of External Studies.

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive psychology*, *63*(3), 107–140. https://doi.org/10.1016/j.cogpsych.2011.06.001

Thompson, V. A., St. Evans, J. B. T., & Campbell, J. I. D. (2013). Matching bias on the selection task: It's fast and feels good. *Thinking & Reasoning, 19*(3–4), 431–452. https://doi.org/10.1080/13546783.2013.820220

Thrän, D., Edenhofer, O., Pahle, M., Schill, E., Steubing, M., & Willts, H. (2023). Minderungsansätze in der Energie- und Kreislaufwirtschaft. In G. P. Brasseur, D. Jacob, & S. Schuck-Zöller (Hrsg.), *Klimawandel in Deutschland: Entwicklung, Folgen, Risiken und Perspektiven* (S. 429–438). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-66696-8

Tiemann, R., & Körbs, C. (2014). Die Fragebogenmethode, ein Klassiker der empirischen didaktischen Forschung. In D. Krüger, I. Parchmann, & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 283–296). Springer Berlin Heidelberg.

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences*, *1*(21), 19–25.

Tulving, E. (2005). Episodic Memory and Autonoesis: Uniquely Human? In H. S. Terrace & J. Metcalfe (Hrsg.), *The Missing Link in Cognition: Origins of self-reflective consciousness* (S. 3–56). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195161564.001.0001

Tversky, A., & Johnson, E. J. (1983). Affect, Generalization, and the Perception of Risk. *Journal of personality and social psychology*, *45*(1), 20.

Universität zu Köln. (2022). *Leitlinien der Universität zu Köln zur guten wissenschaftlichen Praxis vom 25. Januar 2022*. https://am.uni-koeln.de/e35075/am_mitteilungen/@e35516/AM_2022-08_Leitlinien-guter-wiss-Praxis_ger.pdf

Urban, D., & Mayerl, J. (2011). *Regressionsanalyse: Theorie, Technik und Anwendung: Lehrbuch* (4., überarb. und erw. Aufl). VS, Verl. für Sozialwissenschaften.

Van Herk, H., Poortinga, Y. H., & Verhallen, T. M. M. (2004). Response Styles in Rating Scales: Evidence of Method Bias in Data From Six EU Countries. *Journal of Cross-Cultural Psychology*, *35*(3), 346–360. https://doi.org/10.1177/0022022104264126

VERBI - Software. Consult. Sozialforschung. GmbH. (2024). *MAXQDA*. MAXQDA. https://www.maxqda.com/de

Vieira, R. M., & Tenreiro-Vieira, C. (2016). Fostering Scientific Literacy and Critical Thinking in Elementary Science Education. *International Journal of Science and Mathematics Education*, *14*(4), 659–680. https://doi.org/10.1007/s10763-014-9605-2

Vieira, R. M., Tenreiro-Vieira, C., & Martins, I. P. (2011). Critical thinking: Conceptual clarification and its importance in science education. *Science Education International*, *22*(1), 43–54.

Vodafone Stiftung Deutschland gGmbH. (2020). *Die Jugend in der Infodemie: Eine repräsentative Befragung zum Umgang junger Menschen in Deutschland mit Falschnachrichten während der Coronakrise*. Abgerufen 09.09.2024, von https://www.vodafone-stiftung.de/die-jugend-in-der-infodemie/

Völkl, K., & Korb, C. (2018). *Deskriptive Statistik: Eine Einführung für Politikwissenschaftlerinnen und Politikwissenschaftler*. Springer VS.

Von Münch, W. (1993). *Einführung in die Halbleitertechnologie*. Vieweg+Teubner Verlag. https://doi.org/10.1007/978-3-322-88970-6

Von Tschirnhaus, M. K. (2022). Fracking als Chance? *Natur und Recht*, *44*(12), 817–825. https://doi.org/10.1007/s10357-022-4101-x

Voss, T. (2017). James S. Coleman: Foundations of Social Theory. In K. Kraemer & F. Brugger (Hrsg.), *Schlüsselwerke der Wirtschaftssoziologie* (S. 213–223). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-08184-3

Weber, A., Büssing, A. G., Jarzyna, R., & Fiebelkorn, F. (2020). Do German Student Biology Teachers Intend to Eat Sustainably? Extending the Theory of Planned Behavior with Nature Relatedness and Environmental Concern. *Sustainability*, *12*(12), 4909. https://doi.org/10.3390/su12124909

Weber, A., & Fiebelkorn, F. (2019). Nachhaltige Ernährung, Naturverbundenheit und Umweltbetroffenheit von angehenden Biologielehrkräften – Eine Anwendung der Theorie des geplanten Verhaltens. *Zeitschrift für Didaktik der Naturwissenschaften*, *25*(1), 181–195. https://doi.org/10.1007/s40573-019-00098-3

Wegner, D. M. (1994). Ironic Processes of Mental Control. *Pychological review*, *101*(1), 34.

Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, *27*(3), 236–247. https://doi.org/10.1016/j.ijresmar.2010.02.004

Weinert, F. E. (2001). *Leistungsmessungen in Schulen* (F. E. Weinert, Hrsg.; 2., unveränd. Aufl.). Beltz.

Weleda. (2024a). *Umsatz der Weleda Gruppe nach Regionen 2023*. Abgerufen 13.08.2024, von https://de.statista.com/statistik/daten/studie/697438/umfrage/umsatz-der-weleda-gruppe-nach-regionen/

Weleda. (2024b). *Umsatz der Weleda Gruppe weltweit in den Jahren 2011 bis 2023*. Abgerufen 13.08.2024, von https://de.statista.com/statistik/daten/studie/697367/umfrage/umsatz-der-weleda-gruppe-weltweit/

Weng, L.-J. (2004). Impact of the Number of Response Categories and Anchor Labels on Coefficient Alpha and Test-Retest Reliability. *Educational and Psychological Measurement*, *64*(6), 956–972. https://doi.org/10.1177/0013164404268674

Wengert, H. G. (2006). Leistungsbeurteilung in der Schule. In G. Bovet & V. Huwendiek (Hrsg.), *Leitfaden Schulpraxis. Pädagogik und Psychologie für den Lehrberuf* (4. Aufl., S. 294–319). Cornelsen Scriptor.

Wiebelitz, K. R., Weyert Goecke, T., Brach, J., & Beer, A.-M. (2009). Use of complementary and alternative medicine in obstetrics. *British Journal of Midwifery*, *17*(3), 169–175. https://doi.org/10.12968/bjom.2009.17.3.40079

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, *1*(6), 80. https://doi.org/10.2307/3001968

Williams. (2008). Internal and External Reasons. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and its Foundations*. Cambridge University Press.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A Model of Dual Attitudes. *Psychological Review*, *107*(1), 101–126. https://doi.org/10.1037/0033-295X.107.1.101

Winch, C. (2003). Die Entwicklung kritischer Rationalität als pragmatische Aufgabe der Erziehung. In *Kritik in der Pädagogik: Bd. Zeitschrift für Pädagogik, Beiheft* (Nummer 46, S. 13–32). Beltz : Weinheim. https://doi.org/10.25656/01:3958

Wissenschaft im Dialog gGmbH (Hrsg.). (2023). *Wissenschaftsbarometer 2023*.

World Health Organization (Hrsg.). (2021). *WHO Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monooxide*. World Health Organization.

Wright, I. (1992). Critical thinking: Curriculum and instructional policy implications. *Journal of Education Policy*, *7*(1), 37–43. https://doi.org/10.1080/0268093920070103

Yacoubian, H. A. (2015). A Framework for Guiding Future Citizens to Think Critically About Nature of Science and Socioscientific Issues. *Canadian Journal of Science, Mathematics and Technology Education*, *15*(3), 248–260. https://doi.org/10.1080/14926156.2015.1051671

Yacoubian, H. A. (2018). Scientific literacy for democratic decision-making. *International Journal of Science Education*, *40*(3), 308–327. https://doi.org/10.1080/09500693.2017.1420266

Yacoubian, H. A. (2020). Teaching Nature of Science Through a Critical Thinking Approach. In W. F. McComas (Hrsg.), *Nature of science in science instruction: Rationales and strategies* (New, updated and revised edition, S. 199–212). Springer.

Zahoransky, R. (2022a). Konventionelle Dampfkraftwerke. In R. Zahoransky (Hrsg.), *Energietechnik: Systeme zur konventionellen und erneuerbaren Energieumwandlung. Kompaktwissen für Studium und Beruf* (9., überarbeitete und ergänzte Auflage, S. 45–102). Springer Vieweg.

Zahoransky, R. (2022b). Überblick, Energiequellen und Energiebedarf. In R. Zahoransky (Hrsg.), *Energietechnik: Systeme zur konventionellen und erneuerbaren Energieumwandlung. Kompaktwissen für Studium und Beruf* (9., überarbeitete und ergänzte Auflage, S. 25–44). Springer Vieweg.

Zeidler, D. L. (2014). Socioscientific issues as a curriculum emphasis. In N. G. Lederman & S. K. Abell (Hrsg.), *Handbook of research on science education Vol. II* (Bd. 2, S. 697–726). Routledge.

Zeidler, D. L., Herman, B. C., & Sadler, T. D. (2019). New directions in socioscientific issues research. *Disciplinary and Interdisciplinary Science Education Research*, *1*(1), 11. https://doi.org/10.1186/s43031-019-0008-7

Zeidler, D. L., & Kahn, S. (2014). *It's debatable!: Using socioscientific issues to develop scientific literacy, K-12*. NSTA Press.

Zeidler, D. L., & Lewis, J. (2003). Unifying Themes in Moral Reasoning on Sociocientific Issues and Discourse. In D. L. Zeidler (Hrsg.), *The Role of Moral Reasoning on Socioscientific Issues and Discourse in Science Education*. Springer Netherlands. https://doi.org/10.1007/1-4020-4996-X

Zeidler, D. L., & Nichols, B. H. (2009). Socioscientific issues: Theory and practice. *Journal of Elementary Science Education*, *21*(2), 49–58. https://doi.org/10.1007/BF03173684

Zeidler, D. L., & Sadler, T. D. (2007). The Role of Moral Reasoning in Argumentation: Conscience, Character, and Care. In S. Erduran & M. P. Jiménez-Aleixandre (Hrsg.), *Argumentation in Science Education* (Bd. 35, S. 201–216). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6670-2_10

Zoom Video Communications, Inc. (2024). *Zoom*. Zoom. https://zoom.us

Zowada, C., & Eilks, I. (2018). Fracking: Ein kontroverses Thema für den fächerübergreifenden Chemieunterricht multimedial umgesetzt. *MNU Journal*, *71*(4), 246–252.

Zowada, C., Gulacar, O., & Eilks, I. (2018). Incorporating a Web-Based Hydraulic Fracturing Module in General Chemistry as a Socio-Scientific Issue That Engages Students. *Journal of Chemical Education*, *95*(4), 553–559. https://doi.org/10.1021/acs.jchemed.7b00613

Anhang

Tabellenverzeichnis des Anhangs 406

A.1 Ergänzendes Material zur Datenerhebung 407

A.1.1 Grundlegende Struktur des finalen Interviewleitfadens 407

A.1.2 Grundlegende Struktur des finalen Fragebogens.....411

A.1.3 Aufgabenstellungen der Systemischen Stellungnahme.....419

A.1.3.1 Systemische Stellungnahme Aufgabenstellung Teil A 419

A.1.3.2 Systemische Stellungnahme Aufgabenstellung Teil B, differente Teile.....424

A.1.3.3 Systemische Stellungnahme Aufgabenstellung Teil C, differente Teile.....425

A.2 Ergänzendes Material zur Datenauswertung 426

A.2.1 Verwendete Transkriptionsregeln 426

A.3 Kodierleitfäden 427

A.3.1 Kodierleitfäden der Studie S3 427

A.3.1.1 Kodierleitfäden S3.I1&2 + Follow-Up.....427

A.3.1.2 Kodierleitfäden S3.F1&2 442

Tabellenverzeichnis des Anhangs

Tabelle A 1: Interviewleitfaden S3.I1&2 - E-Mobilität vs. Brennstoffzellenantrieb (B) 407

Tabelle A 2: Verwendete modifizierte Transkriptionsregeln nach Dresing uns Pehl (2018, S. 22) sowie Kuckartz (2016, S. 167) 426

Tabelle A 4: Kodierleitfaden S3.I1&2 + FU – Selbsteinschätzung Fachwissen (NW-SF) 427

Tabelle A 5: Kodierleitfaden S3.I1&2 + FU – Fachwissenschaftliche Bewertung (NW-FB) 428

Tabelle A 6: Kodierleitfaden S3.I1&2 + FU – Aspekte naturwissenschaftlichen Wissens (NW-A) 429

Tabelle A 7: Kodierleitfaden S3.I1&2 + FU – Verständnis Selbstreflexion (SR-V) 431

Tabelle A 8: Kodierleitfaden S3.I1&2 + FU – Aspekte der Selbstreflexion (SR-A) 433

Tabelle A 9: Kodierleitfaden S3.I1&2 + FU – Aspekte der Epistemologie (EP-A) 436

Tabelle A 10: Kodierleitfaden S3.I1&2 + FU – Verständnis Kritisches Denken (KD-V) 439

Tabelle A 11: Kodierleitfaden S3.I1&2 + FU – Fähigkeit zur Bewertung (FB) 441

Tabelle A 12: Kodierleitfaden S3.F1&2 – Verständnis Selbstreflexion (VS) 442

Tabelle A 13: Kodierleitfaden S3.F1&2 – Verständnis Kritisches Denken (VK) 445

Tabelle A 14: Kodierleitfaden S3.F1&2 – Beeinflussende Faktoren bei der Bewertung (BF) 447

A.1 Ergänzendes Material zur Datenerhebung

A.1.1 Grundlegende Struktur des finalen Interviewleitfadens

Tabelle A 1: Interviewleitfaden S3.I1&2 - E-Mobilität vs. Brennstoffzellenantrieb (B)

Phase des Interviews	**Aktion/Frage**	**Anmerkungen/ Zusätze**
Generelle Anmerkungen	❖ Um Nähe zu den Proband*innen zu erzeugen, wird das **„Du" als Anrede** gewählt. ❖ Ziel ist eine **ungezwungene Gesprächsatmosphäre**. ❖ **Flexibel bleiben** in der Gesprächsführung. ❖ Fachwörter oder Beschreibungen **genauer erläutern lassen**, um evtl. Fassadenwissen aufzudecken. ❖ Mit **Ruhe** nachfragen und bedacht formulieren **ohne Druck** zu suggerieren und soziale Erwünschtheit zu provozieren. ❖ **Denkpausen tolerieren**. Bei längeren Pausen können die Denkprozesse erfragt werden.	
Einleitung und Vorstellung	➢ Begrüßung/**Lockerer Einstieg** in das Gespräch „Wie geht es dir heute? /Was machst du gerne in deiner Freizeit?" ➢ **Rahmen des Interviews** erläutern. „Dieses Interview findet im Rahmen meiner Dissertation in der Chemiedidaktik an der Universität zu Köln statt. Ich interessiere mich für deine Strategien bei der Meinungsbildung zu gesellschaftsrelevanten Themen. Dieses Interview ermöglicht mir tiefergehende Informationen zu gewinnen, was dir bei deinen Entscheidungen besonders wichtig ist." ➢ **Angst** vor dem Interviewsetting **nehmen.** „Hast du schon an Interviews teilgenommen?" „Es ist keine Prüfung und es gibt kein Richtig oder Falsch!" ➢ Aufklären über den **Datenschutz** (Pseudonymisierung), kurze Abfrage des **Einverständnisses** und gemeinsame **Zuordnung zum Code**. o Ersten zwei Buchstaben der Straße in der Du aufgewachsen bist. o Die zwei Zahlen deines Geburtstages.	Entspannte Gesprächssituation aufbauen, Druck reduzieren. Forschungsthema kurz vorstellen. Nur so viel wie nötig und wenig wie möglich im Vorfeld anreißen, um Priming-Effekte zu vermeiden. Suggestivfragen vermeiden.

	○ Die letzten zwei Buchstaben des Vornamens des Vaters. ➢ Auf das **Recht an den eigenen Daten** hinweisen. ➢ Bitten, nicht mit den **Kommiliton*innen** über den Inhalt des Interviews zu reden, um **Priming** zu **verhindern**. ➢ Interviewablauf erläutern. „Ich werde mich sehr **im Hintergrund halten** und dir zuhören. Am besten wäre es, wenn du immer so viel wie möglich **von dir selbst aus erzählst**, da **deine Ansichten für mich von Interesse** sind." „Sollten **organisatorische Fragen** im Laufe des Interviews aufkommen, beantworte ich sie dir **gerne nach dem Interview**. **Inhaltliche Fragen** können **jederzeit** gestellt werden. Solltest du bis hierher oder zum Ablauf noch Fragen haben, kannst du diese jetzt gerne stellen!"	Datenschutz und Ablauf
Einstieg	◊ **Wie informierst du dich über das aktuelle Tagesgeschehen?** ○ (Nachfrage zu den genannten Wegen oder Medien.)	Hinleitung in Richtung gesellschaftsrelevante Themen, seichter Einstieg, Icebreaker-Frage.
Problemaufriss; Freie Formulierung relevanter Aspekte	◊ Gehen wir von folgendem Setting aus: Du möchtest dich zu einem **Diskurs positionieren**, beispielsweise zur politischen Teilhabe im Rahmen einer Wahl. Es stehen dabei **zwei Positionen einander gegenüber**: Die eine Position steht für die Subventionierung von **Elektrofahrzeugen**, die andere Position für die Subventionierung von **Brennstoffzellenfahrzeugen**. Könntest du jetzt schon sagen, **zu welcher Position** du tendieren würdest? ◊ **Denk für mich ab hier bitte laut: Was ist für dich hier relevant, um eine Entscheidung treffen zu können?**	Erfragen der Gedankengänge bei einer gesellschaftsrelevanten Entscheidung mit naturwissenschaftlichem Bezug ohne klare Lösung.
Formulierung von relevanten Fragen ohne Vorgabe; **Ergründung des Vorgehens im Bewertungskontext**	◊ **Welche Fragen stellst du dir, wenn du vor so eine Entscheidung gestellt wirst?** ○ Bitte sprich deinen **Gedankengang offen** aus, so als würdest du laut denken. ○ Was sind für dich hier **relevante Aspekte**? ○ Welche Fragen hättest du gerne beantwortet, um diese Entscheidung treffen zu können?	<u>Mögliche Formulierungen und Begriffe:</u> Weiß ich genug über das Thema? Was sagt Partei A zu Thema B? Wie wichtig ist mir Thema Y? *Lautes Denken seitens der Proband*innen.* *Formulierung relevanter Aspekte.*

Beschreibung des Vorgehens bei der Meinungsbildung	◊ **Wie würdest du vorgehen, wenn du dich zu dem Kontext positionieren müsstest?** o (Wird ein konkretes Vorgehen genannt, sind zusätzliche Nachfragen möglich.) o Was ist die bei diesem Vorgehen besonders wichtig? o Was ist dir bei Quellen besonders wichtig? o Woran **erkennst du** für dich, **ob eine Quelle verlässlich** ist?	*Formulierung von relevanten Fragen.* *Erläutern des Vorgehens zur Meinungsbildung.*
Fragen nach wissensbezogenen Aspekten	◊ **Reicht dir dein aktuelles Wissen aus, um dir eine Meinung zu dem Thema zu bilden?** ◊ **Wie viel weißt du denn bereits über das Thema?** ◊ **Woher nimmst du dieses Wissen?** o Inwieweit hinterfragst du dein Vorwissen aktiv, bevor du dir eine Meinung über ein Thema bildest? o Fallen dir dann auch Wissenslücken auf? o Wie stark/wie sehr reflektierst du dein Vorwissen und etwaige Lücken, bevor du eine Entscheidung triffst?	*Welches Vorwissen wird herangezogen? Ist es umfassend genug?* <u>Mögliche Antworten:</u> Nein, ich benötige mehr Information. *Ja, ich weiß schon XY darüber und deshalb…*
Fragen nach selbstreflexiven Aspekten	◊ **Wie stehst du bisher zu diesem Thema?** ◊ **Was glaubst du, prägt deine Meinung besonders?** o Reflektiere dich doch einmal selbst. Welche Fragen kommen dabei eventuell auf, die du beantwortet haben wollen würdest oder die du dich selbst fragst? o Inwieweit hinterfragst du, was deine Meinung noch beeinflussen könnte, bevor du dir eine Meinung bildest? ◊ **Was glaubst du, könnte deine Meinung noch beeinflussen?** ◊ **Wie stark reflektierst du den Einfluss anderer Prägungen auf deine Meinungen, bevor du eine Entscheidung triffst (soziales Umfeld, Werbung, Politik, Schule, etc.)?** o Gibt es Dinge, die dich **besonders stark** beeinflussen? o Wie viel Einfluss hat diese Vormeinung auf deine Entscheidung?	Anstoßen eines Reflexionsprozesses *Welche Vorannahmen und Einflüsse werden genannt?* <u>Mögliche Prägungen:</u> Medien, soziales Umfeld, Studium, vorhandene politische Einstellungen…

Frage nach der Selbsteinschätzung und dem Verständnis der Begrifflichkeiten *kritisch* und *reflektiert*	◊ **Würdest du dich selbst als kritisch bezeichnen?** o Falls ja, bitte führe mit Beispielen aus, was du darunter verstehst und wo es erkennbar wird. (Möglichst an diesem Beispiel.) ◊ **Was verstehst du unter dem Begriff „kritisch“?** ◊ **Würdest du dich selbst als reflektiert bezeichnen?** o Falls ja, bitte führe mit Beispielen aus, was du darunter verstehst und wo es erkennbar wird. (Möglichst an diesem Beispiel.) ◊ **Was verstehst du unter dem Begriff „reflektiert“?**	Ermöglicht Rückschlüsse auf das eigene Verständnis der Begrifflichkeit *kritisch* und die Selbsteinschätzung. Beispiele machen deutlich, ob ein klares und konkretes Bild vorliegt. Explizite Einschätzung der eigenen Reflexionsfähigkeit und Selbsteinschätzung. Ermöglicht Rückschlüsse auf das Verständnis der Begrifflichkeit *reflektiert*.
Abschluss	◊ Gibt es noch etwas, das du ergänzen möchtest? ◊ Hast du Fragen?	

A.1.2 Grundlegende Struktur des finalen Fragebogens

Aufbau des finalen Fragebogens, queXML-Export aus der LimeSurvey-Umfrage zur Erhebung S5.F1

Herzlich Willkommen zur digitalen Umfrage im Rahmen der Dissertation von Luc Albrecht.

Vielen Dank, dass Sie sich dazu entschieden haben, mir bei meinem Foschungsvorhaben zu helfen!

Im Rahmen meiner Dissertation an der Universität zu Köln beschäftige ich mich mit der Selbsteinschätzung von angehenden Chemielehrer*innen bezüglich ihrer Bewertungsfähigkeit.

Viel mehr sollte an dieser Stelle nicht genannt werden, da ich eine Vorbeeinflussung meinerseits verhindern möchte.

Nur so viel:

Bitte machen Sie Ihre Angaben so ehrlich wie möglich.

Ihre personenbezogenen Daten und Antworten werden pseudonymisiert über einen Code erhoben und absolut vertraulich behandelt! Ihre Daten und Antworten werden lediglich für die Auswertung im Rahmen der Dissertation verwendet und nicht darüber hinaus weitergegeben.

Der Fragebogen dauert etwa 25 - 35 Minuten.

Und nun freue ich mich auf Ihre Antworten und wünsche Ihnen viel Spaß!

Mit dankbaren Grüßen,

Luc Albrecht

Bei weiteren Fragen kontaktieren Sie mich gerne unter:

l.albrecht@uni-koeln.de

Zunächst ein paar Fragen zu Ihrer Person und Ihrem bisherigen Studienverlauf:

A1. **Erstellung Ihres persönlichen Codes**

Zum Schutze Ihrer persönlichen Daten werden diese mittels eines Codes pseudonymisiert. Dies ermöglicht mir die Verknüpfung Ihrer Antworten mit Ihrer Person für mögliche Nachfragen und Interviews, erschwert aber Dritten die Zuordnung und Identifikation.

Ihr individueller Code setzt sich wie folgt zusammen:

Die ersten zwei Buchstaben der Straße in der Sie aufgewachsen sind. Bsp.: Kanalstraße = KA
Die zwei Zahlen Ihres Geburtstages (nicht des Monats). Bsp.: Geburtstag am 26.03.1986 = 26
Die letzten zwei Buchstaben des Vornamens Ihres Vaters. Bsp.: Bernd = ND

Der komplette Code wäre dann : KA26ND

Tragen Sie nun bitte Ihren persönlichen Code in das untenstehende Feld ein.

A2. **Welchen Lehramtsstudiengang belegen Sie derzeit?**

- ☐ Lehramt für Haupt-, Real- und Gesamtschulen
- ☐ Lehramt für Gymnasien und Gesamtschulen
- ☐ Lehramt für Berufskollegs
- ☐ Lehramt für Grundschule
- ☐ Lehramt für Sonderpädagogik
- ☐ Sonstiges

Sonstiges

A3. **Welches sind ihre weiteren Fächer (neben Chemie)?**

1. ☐☐☐☐☐☐☐☐☐☐
2. ☐☐☐☐☐☐☐☐☐☐
3. ☐☐☐☐☐☐☐☐☐☐

A4. **In welchem Abschnitt Ihres Studiums befinden Sie sich gerade?**

- ☐ Bachelor
- ☐ Master
- ☐ Promotion

A5.

Im wievielten Semester sind Sie aktuell?

(Gemeint sind die Semester im derzeitigen Studiengang, also im Bachelor oder Master, nicht die Gesamtsemesteranzahl!

Sollten Sie Ihr Studium vor Kurzem abgeschlossen haben, tragen Sie bitte "*abgeschlossen*" ein)

Folgendes Setting definiert die Ausgangslage: Sie werden für eine politische Umfrage befragt, die für die zukünftige Mobilitätsausrichtung des Landes relevant ist. Aus Klimaschutzgründen wird der Verbrenner abgeschafft und es stehen zwei alternative Antriebsmöglichkeiten zur Diskussion: die Elektromobilität und der Wasserstoffantrieb. Die Frage, die nun im Raum steht, ist folgende:

Sollte der Ausbau und die Förderung von Elektrofahrzeugen verstärkt werden oder sollte der Ausbau und die Förderung von Brennstoffzellenautos verstärkt werden,

um den Einsatz von fossilen Brennstoffen im Verkehrssektor zu reduzieren?

B1. **Wie würden Sie sich entscheiden und warum? Bitte führen Sie alle für Sie relevanten Aspekte an, um sich klar positionieren zu können.**

Sie müssen nicht bereits alle Informationen, die für Sie relevant wären, vorliegen haben.

In diesem Falle stellen Sie gerne Fragen nach den für Sie relevanten Aspekten.

B2. **Welche Aspekte wären aus fachlicher bzw. naturwissenschaftlicher Sicht für Sie relevant, um sich klar positionieren zu können?**

Sie müssen nicht bereits alle Informationen, die für Sie relevant wären, vorliegen haben.

In diesem Falle stellen Sie gerne Fragen nach den für Sie relevanten Aspekten.

B3. **Welche Aspekte wären mit Blick auf sich selbst für Sie relevant, um sich klar positionieren zu können?**

Sie müssen nicht bereits alle Informationen, die für Sie relevant wären, vorliegen haben.

In diesem Falle stellen Sie gerne Fragen nach den für Sie relevanten Aspekten.

B4. **Welche Aspekte sind für Sie relevant, wenn Sie Quellen hinsichtlich ihrer Glaubwürdigkeit einschätzen müssen?**

Sie müssen nicht bereits alle Informationen, die für Sie relevant wären, vorliegen haben.

In diesem Falle stellen Sie gerne Fragen nach den für Sie relevanten Aspekten.

Es folgen ein paar kurze Fragen zu Ihrem Verständnis. Dies ist kein Wissenstest, sondern soll nur Ihr individuelles Verständnis widerspiegeln.

C1. **Was verstehen Sie unter dem Begriff "Selbstreflexion"?**

C2. **Was verstehen Sie unter "Kritischem Denken"?**

C3. **Was verstehen Sie unter "kognitiven Verzerrungen"?**

C4. **Bitte geben Sie Ihre Antwort für die folgende Frage in das untenstehende Feld ein:**

Ein Schläger und ein Ball kosten zusammen 1,10 €. Der Schläger kostet 1,00 € mehr als der Ball.

Wie viel kostet der Ball?

Euro: ☐☐☐☐☐☐☐☐☐☐

C5. **Was glauben Sie, was beeinflusst Sie besonders, wenn Sie im Alltag Informationen bewerten, die einen naturwissenschaftlichen Bezug haben?**

Bitte beantworten Sie die Fragen ehrlich. Denken Sie daran, es ist keine Prüfung und es findet keine Bewertung statt.

D1. **Geben Sie bitte für jede der kommenden Fragen Ihre Antwort in der nebenstehenden Skala an.**

1= nie/gar nicht ; 5=Immer/sehr ausführlich

	1 (nie / gar nicht)	2 (selten / kaum)	3 (manchmal / mäßig umfassend)	4 (häufig / umfassend	5 (immer / sehr umfassend)
Wie oft machen Sie sich Ihre eigenen Gedanken und Gefühle bewusst, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	☐	☐	☐	☐	☐
Inwieweit setzen Sie sich mit Ihren eigenen Moral- und Wertvorstellungen auseinander, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	☐	☐	☐	☐	☐
Wie häufig denken Sie über Ihre persönlichen Erfahrungen und Hintergründe nach, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	☐	☐	☐	☐	☐
Inwiefern reflektieren Sie Ihre Interessen und Motivationen, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	☐	☐	☐	☐	☐
Wie häufig hinterfragen Sie Ihre eigenen Vorurteile und Annahmen, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?	☐	☐	☐	☐	☐

Im Folgenden werden Ihnen eine Reihe an Aussagen genannt, zu denen Sie bitte den Grad Ihrer Zustimmung angeben.

Von 1 (ich stimme überhaupt nicht zu) hin zu 7 (ich stimme voll und ganz zu).

E1. **Geben Sie für die folgenden Aussagen bitte den Grad Ihrer Zustimmung an:**

	ich stimme überhaupt nicht zu1	2	3	4	5	6	ich stimme voll und ganz zu7
Ich betrachte häufig meine Gefühlswelt	☐	☐	☐	☐	☐	☐	☐
Ich nehme mir häufig Zeit, um über meine Gedanken zu reflektieren	☐	☐	☐	☐	☐	☐	☐
Ich denke oft über die Art und Weise nach, wie ich mich in bestimmten Situationen fühle	☐	☐	☐	☐	☐	☐	☐

	ich stimme überhaupt nicht zu1	2	3	4	5	6	ich stimme voll und ganz zu7
Es ist mir wichtig, die Dinge zu bewerten, die ich tue	☐	☐	☐	☐	☐	☐	☐
Ich bin sehr daran interessiert, nachzuvollziehen, wie ich denke	☐	☐	☐	☐	☐	☐	☐
Es ist mir sehr wichtig, zu verstehen, was meine Gefühle bedeuten	☐	☐	☐	☐	☐	☐	☐

E2. **Bitte geben Sie Ihre Antwort für die folgende Frage in das untenstehende Feld ein:**

Wenn 5 Maschinen 5 Minuten benötigen, um 5 Geräte herzustellen, wie lange dauert es mit 100 Maschinen 100 Geräte herzustellen?

Minuten: ☐☐☐☐☐☐☐☐☐☐

E3. **Geben Sie für die folgenden Aussagen bitte den Grad Ihrer Zustimmung an:**

	ich stimme überhaupt nicht zu1	2	3	4	5	6	ich stimme voll und ganz zu7
Ich bin oft verwundert über die Art und Weise, wie ich wirklich über Dinge denke	☐	☐	☐	☐	☐	☐	☐
Ich bin mir meiner Gefühle bewusst, kann diese aber oft nicht klar benennen	☐	☐	☐	☐	☐	☐	☐
Mein Verhalten erstaunt mich oft	☐	☐	☐	☐	☐	☐	☐
Das Nachdenken über meine Gedanken bringt mich durcheinander	☐	☐	☐	☐	☐	☐	☐
Ich finde es oft schwierig meine Gefühlswelt zu verstehen	☐	☐	☐	☐	☐	☐	☐
Ich weiß normalerweise, warum ich so fühle, wie ich fühle	☐	☐	☐	☐	☐	☐	☐

Bitte beantworten Sie die Fragen ehrlich. Denken Sie daran, es ist keine Prüfung und es findet keine Bewertung statt.

F1. **Welche Faktoren haben Ihrer Meinung nach einen Einfluss auf Ihre Entscheidungen, wenn Sie ein Thema mit naturwissenschaftlichem Bezug bewerten sollen?**

☐

F2. **Bitte geben Sie die Antwort für die folgende Frage in das untenstehende Feld ein:**

In einem See ist ein Feld mit Seerosen. Jeden Tag verdoppelt sich die Größe des Feldes.

Wenn die Seerosen 48 Tage brauchen, um den gesamten See zu bedecken, wie lange dauert es, bis sie die Hälfte des Sees bedecken?

Tage: ☐☐☐☐☐☐☐☐☐☐

F3. **Inwieweit glauben Sie, dass Sie in der Lage sind, vornehmlich auf fachlicher Ebene zu urteilen?**

	1 (überhaupt nicht)	2 (in geringem Maße)	3 (mäßig)	4 (vorwiegend)	5 (vollständig)
Geben Sie auf der Skala Ihre Antwort an:	☐	☐	☐	☐	☐

F4. **Wie sehr hinterfragen Sie folgende Faktoren, wenn Sie sich im Alltag zu Themen eine Meinung bilden, die einen naturwissenschaftlichen Bezug haben?**

	1 (gar nicht)	2 (wenig)	3(mäßig)	4 (ziemlich umfassend)	5 (sehr umfassend)
Eigene Persönlichkeit	☐	☐	☐	☐	☐
Eigene Stimmung und Emotionen	☐	☐	☐	☐	☐
Eigene Werte und Moralvorstellungen	☐	☐	☐	☐	☐
Eigene Grundeinstellungen	☐	☐	☐	☐	☐
Eigene Bildung	☐	☐	☐	☐	☐
Eigenes Alter	☐	☐	☐	☐	☐
Eigenes Geschlecht	☐	☐	☐	☐	☐
Eigene Soziale Schicht	☐	☐	☐	☐	☐
Eigene Religion	☐	☐	☐	☐	☐
Eigene Ethnie	☐	☐	☐	☐	☐
Eigene Kultur	☐	☐	☐	☐	☐
Eigenes Wissen	☐	☐	☐	☐	☐
Medialer Kontext	☐	☐	☐	☐	☐
eigene politische Einstellung	☐	☐	☐	☐	☐
eigene Erfahrungen und Erinnerungen	☐	☐	☐	☐	☐
Selbstkonzept und Selbstwertgefühl	☐	☐	☐	☐	☐
Soziale Normen und Erwartungen	☐	☐	☐	☐	☐
Eigene Motivation und Bedürfnisse	☐	☐	☐	☐	☐
Verfügbarkeit von Alternativen	☐	☐	☐	☐	☐
Eigene Intuition und Bauchgefühl	☐	☐	☐	☐	☐
Gruppendynamik und Konformitätsdruck	☐	☐	☐	☐	☐
Erwartungen und Vorurteile	☐	☐	☐	☐	☐
Framing und Präsentation von Informationen	☐	☐	☐	☐	☐
Selektive Wahrnehmung und Bestätigungsfehler	☐	☐	☐	☐	☐
Übergeneralisierung und Kategorisierung	☐	☐	☐	☐	☐
Komplexität der Information	☐	☐	☐	☐	☐

Bitte beantworten Sie die Fragen ehrlich. Denken Sie daran, es ist keine Prüfung und es findet keine Bewertung statt.

G1. **Geben Sie Ihre Antwort für die folgenden Fragen bitte in der nebenstehenden Skala an.**

	1 (nie / gar nicht)	2 (selten / schwach)	3 (manchmal / mäßig)	4 (häufig / stark)	5(immer / sehr stark)
Wie stark schätzen sie den Einfluss der vorhin genannten Faktoren (Bauchgefühl, Stimmung usw.) im Vergleich zum Fachwissen, wenn Sie ein Thema mit naturwissenschaftlichem Bezug bewerten sollen?	☐	☐	☐	☐	☐
Inwieweit wägen Sie diese Faktoren (Bauchgefühl. Stimmung usw.) bewusst gegen Ihr Fachwissen ab, wenn sie ein Thema mit naturwissenschaftlichem Bezug bewerten sollen?	☐	☐	☐	☐	☐

Hervorragend, Sie haben es geschafft und die Umfrage erfolgreich beendet!

Vielen Dank für Ihre Teilnahme und damit Mithilfe an meiner Dissertation.

Sollten Sie noch etwaige Fragen haben, kontaktieren Sie mich gerne unter:

l.albrecht@uni-koeln.de

A.1.3 Aufgabenstellungen der Systemischen Stellungnahme

A.1.3.1 Systemische Stellungnahme Aufgabenstellung Teil A

1

Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP)

Dozent: L. Albrecht
28.06.2023
Dialogisches Selbstreflexionstraining

45 Min.

Aufgabe 1: Einzelarbeit

Analysieren Sie einen der untenstehenden Kontexte zur Beantwortung der dazugehörigen Fragestellung.

Dokumentieren Sie Ihre Ergebnisse und **bewerten** Sie die Herkünfte Ihrer Informationen.

Kontext 1: Fracking (Hydraulic Fracturing)

Fragestellung: Sollte die Praxis des Frackings, also die hydraulische Frakturierung von Gesteinsschichten zur Gewinnung von Erdgas oder Erdöl, fortgesetzt werden, um Energievorräte zu erschließen und die wirtschaftliche Entwicklung zu fördern, oder sollten wir aufgrund möglicher Umweltauswirkungen und Risiken den Einsatz von Fracking einschränken oder verbieten?

Kontext 2: Plastik

Fragestellung: Sollten Einwegplastikprodukte, wie Plastiktüten und Einwegflaschen, stärker reguliert oder verboten werden, um die Auswirkungen auf die Umwelt durch Plastikmüll zu reduzieren, oder sollten wir den Schwerpunkt eher auf verbessertes Recycling und die Entwicklung biologisch abbaubarer Kunststoffe legen, um die Vorteile von Plastikprodukten beizubehalten?

Kontext 3: Nanotechnologie

Fragestellung: Sollte die Entwicklung und Anwendung von Nanotechnologie strenger reguliert werden, um potenzielle Risiken für die Gesundheit und Umwelt zu minimieren, oder sollten wir die Chancen und Potenziale der Nanotechnologie nutzen und die Regulierung flexibel gestalten, um Innovationen und wirtschaftlichen Fortschritt zu ermöglichen?

Kontext 4: Chemie in der Landwirtschaft

Fragestellung: Sollten chemische Pestizide und Düngemittel in der Landwirtschaft weiterhin weit verbreitet eingesetzt werden, um Ernteerträge zu maximieren und die Welternährungssicherheit zu gewährleisten, oder sollten wir verstärkt auf alternative Methoden, wie biologischen Anbau und nachhaltige Landwirtschaft, setzen, um Umweltauswirkungen und potenzielle Gesundheitsrisiken zu minimieren?

Kontext 5: Atomenergie

Fragestellung: Sollte Atomenergie als eine bedeutende Energiequelle in der Zukunft weiter ausgebaut werden, um den steigenden Energiebedarf zu decken und den Klimawandel zu bekämpfen, oder sollten wir verstärkt auf erneuerbare Energien setzen und Atomenergie aufgrund ihrer potenziellen Risiken und Auswirkungen einschränken oder sogar auslaufen lassen?

1

Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP)

Dozent: L. Albrecht
28.06.2023
Dialogisches Selbstreflexionstraining

Aufgabe 2: Partnerarbeit (Tandem)

45 Min.

Finden Sie sich in Zweiergruppen (Tandems) zusammen. Je einmal Nummer 1 und 2 (s. oben rechts).

Geben Sie **zu Beginn jeweils ein kurzes Statement** ab, welchen Kontext Sie gewählt haben und wie Sie sich positioniert haben.

Stellen Sie sich dann **gegenseitig und abwechselnd** die Fragen aus Ihrem **Impulsleitfaden** und ***reflektieren* Sie sich** anhand der Fragen selbst.

Dokumentieren Sie Ihre Antworten und Ihre gewonnenen Einsichten.

Anmerkungen:

Der/die **Fragenstellende** hat die Aufgabe den/die Befragte*n bei ihrer/seiner **Selbstreflexion zu unterstützen, ohne jedoch zu bewerten**. Dabei hat der/die Fragende immer die Möglichkeit, weiter auf die Antwort einzugehen und die Frage so zu vertiefen.

Der/die **Befragte** sollte versuchen die Fragen so ehrlich wie möglich zu beantworten und über sich selbst und seine Beweggründe nachzudenken. Seien Sie offen für weiterführende Fragen und Impulse.

Merke: Diese Situation sollte ein bewertungsfreier Raum sein. Es gibt hier keine richtigen oder falschen Antworten! Achten Sie auf einen wertschätzenden Umgang miteinander.

Zusatzaufgabe:

Sollten Sie schon früher fertig sein, ***entwickeln*** Sie gemeinsam weitere Fragestellungen, die Ihnen helfen können, sich selbst besser zu reflektieren und kognitiven Verzerrungen vorzubeugen.

Dokumentieren Sie diese ebenfalls schriftlich.

1

Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP)

Dozent: L. Albrecht
28.06.2023
Dialogisches Selbstreflexionstraining

Impulsleitfaden

1. Gibt es eine Beziehung zwischen Deiner Persönlichkeit und Deinen Überzeugungen zu diesem Thema?
2. Wie fühlst Du dich, wenn Du über das Thema nachdenkst?
3. Hast Du stereotype Vorstellungen von Menschen, die eine andere Meinung vertreten?
4. Fühlst Du dich ausreichend informiert, um Dir eine fundierte Meinung bilden zu können?
5. Hat Deine Generation eine spezifische Einstellung zu diesem Thema?
6. Inwiefern können Geschlechterrollen und -erwartungen Deine Meinung zu diesem Thema beeinflussen?
7. Gibt es bestimmte religiöse Überzeugungen, die Deine Meinung beeinflussen?
8. Wie beeinflussen Deine Freunde Deine Meinungsbildung?
9. Wie beeinflusst Deine politische Einstellung Deine Meinung zum Thema?
10. Wie beeinflussen gesellschaftliche Normen und Erwartungen Deine Meinung zum Thema?
11. Fühlst Du dich unter Druck gesetzt, eine bestimmte Meinung zu vertreten?
12. Hast Du alternative Standpunkte in Betracht gezogen, bevor Du Deine Meinung abschließend gebildet hast?
13. Wie beeinflusst die Art und Weise, wie Informationen präsentiert werden, Deine Meinung zum Thema?
14. Wie beeinflussen Deine eigenen Lebenserfahrungen Deine Meinung zu diesem Thema?

90 Min.

Aufgabe 3: Gruppenarbeit (4-6 Personen)

Finden Sie sich in Gruppen von vier bis sechs Personen zusammen.

Gehen Sie auf https://www.rechner.club/zufall/zufallsauswahl-generieren und lassen Sie sich für Ihre Gruppe eine Zufallsreihenfolge generieren. Dies ist die Präsentationsreihenfolge Ihrer Ergebnisse.

Jede Präsentation soll einen Umfang von **5 Minuten** haben, gefolgt von einer kurzen Diskussionsphase von ebenfalls **5 Minuten**.

Als Präsentierende/r:

Nehmen Sie vor der Gruppe ***Stellung*** zu Ihrem gewählten Kontext. Diese Stellungnahme soll sowohl Ihre Ergebnisse der Sachanalyse (Aufgabe 1) als auch die gewonnenen Einsichten aus der Selbstreflexion (Aufgabe 2) beinhalten und gegeneinander abwägen.

Diskutieren Sie, wie sich Ihre Positionierung zur Fragestellung entwickelt hat und inwieweit Ihre Selbstreflexion Veränderungen hervorgerufen oder Fragen aufgeworfen hat.

Bewerten Sie dabei auch, wie groß die Einflüsse auf affektiver Ebene für Sie waren und welche Schwierigkeiten Ihnen bei Ihrer Positionierung begegnet sind.

Als Zuhörende/r in der Diskussionsphase:

Versuchen Sie die Beweggründe, die zu einer Positionierung der/des Vortragenden geführt haben, zu verstehen. **Stellen Sie Fragen**, welche sowohl Ihnen als auch der/dem Präsentierenden helfen relevante Faktoren für die Positionierungen herauszustellen und zu reflektieren.

Merke: Diese Situation sollte ein bewertungsfreier Raum sein. Es gibt hier keine richtigen oder falschen Antworten! Achten Sie auf einen wertschätzenden Umgang miteinander.

Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP)

Dozent: L. Albrecht
28.06.2023
Dialogisches Selbstreflexionstraining

Aufgabe 4: Einzelarbeit

30 Min.

Reflektieren Sie Ihre gemachten Erfahrungen und Ihre Entscheidungsfindungsprozesse.

Diskutieren Sie die Herausforderungen, die bei einer Positionierung auftreten können, und gehen Sie dabei auf Ihre in den letzten Aufgaben gewonnenen Erkenntnisse ein.

Erläutern Sie, welche kognitiven Verzerrungen für Sie besonders spürbar waren und welche Strategie Sie angewendet haben, um deren Einfluss zu reduzieren.

Dokumentieren Sie Ihre Ergebnisse und laden Sie diese im ILIAS-Ordner der heutigen Sitzung hoch.

Benennen Sie die Datei nach folgendem Schema:

Nachname_DST_Datum Beispiel: Müller_DST_28.06.2023

A.1.3.2 Systemische Stellungnahme Aufgabenstellung Teil B, differente Teile

2

Seminar zu fachbezogenen Lern- und Kommunikationsprozessen (B5 HR/GG B3 SP)

Dozent: L. Albrecht
28.06.2023
Dialogisches Selbstreflexionstraining

Impulsleitfaden

1. Bist Du tendenziell eher pessimistisch oder optimistisch eingestellt, wenn es um das Thema geht?
2. Wie beeinflusst Deine Stimmung die Art und Weise, wie Du über dieses Thema denkst und sprichst?
3. Wie stehst Du generell zu wissenschaftlichen Erkenntnissen und Fakten?
4. Bist Du bereit, Deine Meinung zu ändern, wenn Du auf neue Informationen stößt?
5. Gibt es eine bestimmte soziale Gruppe, die Deine Meinung stärker beeinflusst als andere?
6. Wie beeinflusst Deine Kultur Deine Meinung zum Thema?
7. Wie haben Deine Erziehung und Deine familiären Hintergründe Deine Meinungsbildung beeinflusst?
8. Wie gehst Du damit um, wenn Du dich von familiären Erwartungen oder Werten beeinflusst fühlst?
9. Wie gehst Du damit um, wenn Du eine abweichende Meinung zu der Deiner Freunde vertrittst?
10. Wie beeinflussen Medienberichte und -kommentare Deine Meinung zum Thema?
11. Fühlst Du dich unsicher in Bezug auf Deine Meinung zu diesem Thema?
12. Gibt es bestimmte persönliche Interessen, die Deine Meinung beeinflussen?
13. Wie beeinflussten Deine Intuition und Dein Bauchgefühl Deine Meinung zum Thema?
14. Hast Du bestimmte Erwartungen oder Vorurteile, die Deine Meinung beeinflussen könnten?
15. Wie gehst Du mit Informationen um, die Deine Meinung infrage stellen?

A.1.3.3 Systemische Stellungnahme Aufgabenstellung Teil C, differente Teile

1

Seminar zu ausgewählten Themengebieten des Chemieunterrichts (M01)

L. Albrecht
30.11.2023
Systemische Stellungnahme

45 Min.

Aufgabe 1: Einzelarbeit

Analysieren Sie einen der untenstehenden Kontexte zur Beantwortung der dazugehörigen Fragestellung.

Dokumentieren Sie Ihre Ergebnisse und **bewerten** Sie die Herkünfte Ihrer Informationen.

Kontext 1: Nanotechnologie

Fragestellung: Sollte die Entwicklung und Anwendung von Nanotechnologie strenger reguliert werden, um potenzielle Risiken für die Gesundheit und Umwelt zu minimieren, oder sollten wir die Chancen und Potenziale der Nanotechnologie nutzen und die Regulierung flexibel gestalten, um Innovationen und wirtschaftlichen Fortschritt zu ermöglichen?

Kontext 2: Chemie in der Landwirtschaft

Fragestellung: Sollten chemische Pestizide und Düngemittel in der Landwirtschaft weiterhin weit verbreitet eingesetzt werden, um Ernteerträge zu maximieren und die Welternährungssicherheit zu gewährleisten, oder sollten wir verstärkt auf alternative Methoden, wie biologischen Anbau und nachhaltige Landwirtschaft, setzen, um Umweltauswirkungen und potenzielle Gesundheitsrisiken zu minimieren?

Kontext 3: Atomenergie

Fragestellung: Sollte Atomenergie als eine bedeutende Energiequelle in der Zukunft weiter ausgebaut werden, um den steigenden Energiebedarf zu decken und den Klimawandel zu bekämpfen, oder sollten wir verstärkt auf erneuerbare Energien setzen und Atomenergie aufgrund ihrer potenziellen Risiken und Auswirkungen einschränken oder sogar auslaufen lassen?

Kontext 4: Elektromobilität vs. Wasserstoffantrieb

Sollte zur Verringerung des Einsatzes fossiler Brennstoffe im Verkehrssektor und zur Bekämpfung des Klimawandels die Entwicklung und Förderung von Elektrofahrzeugen, die bereits über eine gute Effizienz und Infrastruktur verfügen, verstärkt werden, oder sollte anderen alternativen Antriebstechnologien, wie z. B. Wasserstoff-Brennstoffzellen, höhere Priorität eingeräumt werden, um die potenziellen Umweltauswirkungen der Batterieherstellung und -entsorgung zu minimieren?

A.2 Ergänzendes Material zur Datenauswertung

A.2.1 Verwendete Transkriptionsregeln

Tabelle A 2: Verwendete modifizierte Transkriptionsregeln nach Dresing uns Pehl (2018, S. 22) sowie Kuckartz (2016, S. 167)

1. Es wird wörtlich transkribiert, also nicht lautsprachlich oder zusammenfassend. Vorhandene Dialekte werden nicht mit transkribiert, sondern möglichst genau in Hochdeutsch übersetzt.
2. Sprache und Interpunktion werden leicht geglättet, d. h. an das Schriftdeutsch angenähert. Zum Beispiel wird aus „Er hatte noch so'n Buch genannt“ → „Er hatte noch so ein Buch genannt". Die Satzform, bestimmte und unbestimmte Artikel etc. werden auch dann beibehalten, wenn sie Fehler enthalten.
3. Deutliche, längere Pausen von drei Sekunden werden durch in Klammern gesetzte Auslassungspunkte (...) markiert. Bei längeren Pausen wird eine Zahl entsprechend der Dauer in Sekunden angegeben.
4. Zustimmende bzw. bestätigende Lautäußerungen der Interviewenden (mhm, aha etc.) werden nicht mit transkribiert, sofern sie den Redefluss der befragten Person nicht unterbrechen.
5. Einwürfe der jeweils anderen Person werden in Klammern gesetzt.
6. Lautäußerungen der befragten Person, die die Aussage unterstützen oder verdeutlichen (etwa Lachen oder Seufzen), werden in Klammern notiert.
7. Absätze der interviewenden Person werden durch ein „I:“, die der befragten Person(en) durch das Kürzel, z.B. “B:", gekennzeichnet. Das Probandenpseudonym zur Zuordnung wird in der Kopfzeile einer jeden Seite aufgeführt.
8. Jeder Sprechbeitrag wird als eigener Absatz transkribiert. Sprecher*innenwechsel wird durch zweimaliges Drücken der Enter-Taste, also einer Leerzeile zwischen den Sprecher*innen, deutlich gemacht, um so die Lesbarkeit zu erhöhen.
9. Störungen werden unter Angabe der Ursache in Klammern notiert, z.B. (Handy klingelt).
10. Nonverbale Aktivitäten und Äußerungen der befragten wie auch der interviewenden Person werden in Klammern notiert, z.B. (lacht), (stöhnt) und Ähnliches.
11. Unverständliche Wörter werden durch (unv.) kenntlich gemacht.
12. Alle Angaben, die einen Rückschluss auf eine befragte Person erlauben, werden anonymisiert.

A.3 Kodierleitfäden

A.3.1 Kodierleitfäden der Studie S3

A.3.1.1 Kodierleitfäden S3.I1&2 + Follow-Up

Tabelle A 3: Kodierleitfaden S3.I1&2 + FU – Selbsteinschätzung Fachwissen (NW-SF)

S3.I1&2 - Naturwissenschaftliches Wissen (NW)					
Wie schätzen Studierende des Chemielehramts ihr Fachwissen in dem gestellten Kontext ein? (NW-SF)					
Kontexteinheit:	**Kodiereinheit:**	**Abstraktionsniveau:**	**Strukturdimension:**		**Auswertungseinheit:**
Das Transkript zum jeweiligen, personen-bezogenen Interview.	Klare, bedeutungstragende Elemente im Interview, wie Satzteile oder Wortzusammenhänge.	Generalisierbare Aussagen zu Faktoren, Kriterien oder Formulierungen, welche aufzeigen, wie das eigene Fachwissen subjektiv eingeschätzt wird. Geringes Abstraktionsniveau.	Explizit genannte Faktoren, Kriterien oder Formulierungen sowie implizite Hinweise, welche aufzeigen, wie das eigene Fachwissen subjektiv eingeschätzt wird		Alle Interviewtranskripte der Studie S3.
Kategorie:		**Definition:**	**Ankerbeispiel:**	**Kordierregel:**	**Kodierung:**
Selbsteinschätzung Fachwissen (SF)	**NW-SF-UZ: Unzureichend (UZ)**	Das **eigene Fachwissen** über den Kontext wird als **unzureichend** bewertet.	"Ja, also wirklich, wirklich nicht viel. Also habe ich mich eigentlich echt kaum mit beschäftigt. Ja, also. Müsste mir man wahrscheinlich echt so auf einem kompletten Laienlevel alles praktisch erklären, weil ich, weil ich da wirklich raus bin." (S3.I2-SS23-B-B3_5_XJ36JE, Pos. 35)	Erfüllt, wenn das eigene Fachwissen als **nicht ausreichend** angesehen wird, um den gegebenen Kontext bewerten zu können.	**N=0**
	NW-SF-GL: Grundlegend (GL)	Das **eigene Fachwissen** über den Kontext wird als **grundlegend** bewertet.	"Also ich will erst mal sagen, dass ich so generell so naturwissenschaftliches Grundverständnis habe und mich auch für solche Themen interessiere." (S2.I1-Wi22_23-A2-B2_3-XJ27BO, Pos. 42)	Erfüllt, wenn das eigene Fachwissen als **grundlegend** angesehen wird, um den gegebenen Kontext bewerten zu können.	**N=1**
	NW-SF-UR: Umfangreich (UR)	Das **eigene Fachwissen** über den Kontext wird als **umfangreich** bewertet.	"Ich würde sagen, ich bin durchaus gut informiert über diese, über beide Themen, also sowohl Wasserstoff, E-Fuels als auch Elektrotechnik, weil ich das persönlich auch sehr interessant finden." (S3.I1-SS23-B-B3_5-WJ32FE, Pos. 38)	Erfüllt, wenn bei der Bewertung des Kontextes das eigene Fachwissen als ausreichend **umfangreich** angesehen wird, um den Kontext **auf wissenschaftlicher Ebene** einzuschätzen.	**N=2**

Tabelle A 4: Kodierleitfaden S3.I1&2 + FU – Fachwissenschaftliche Bewertung (NW-FB)

S3.I1&2 - Naturwissenschaftliches Wissen (NW)				
Inwieweit findet bei Studierenden des Chemielehramts Fachwissen Einzug in die Bewertung des Kontextes? (NW-FB)				
Kontext-einheit:	**Kodiereinheit:**	**Abstraktionsniveau:**	**Strukturdimension:**	**Auswertungsein-heit:**
Das Tran-skript zum jeweiligen, personen-bezogenen Interview.	Klare, bedeu-tungstragende Elemente im In-terview, wie Satz-teile oder Wortzu-sammenhänge.	Generalisierbare Aussagen zu Faktoren, Kriterien oder For-mulierungen, welche aufzeigen, inwieweit Wissen zur Bewer-tung des Kontextes herangezo-gen wird. Mittleres Abstrakti-onsniveau.	Explizit genannte Faktoren, Krite-rien oder Formulierungen sowie implizite Hinweise, welche aufzei-gen, welche Art von Wissen bei der Bewertung des Kontextes hin-zugezogen wurde.	Alle Inter-viewtranskripte der Studie S3.

Kategorie:		**Definition:**	**Ankerbeispiel:**	**Kordierregel:**	**Niveau:**
Fachwissenschaftliche Bewertung (FB)	**NW-FB-NV: Nicht vorhan-den (NV)**	Es findet **keine Bewertung** auf Grundlage **natur-wissenschaftli-chen Fachwis-sens** statt.		Erfüllt, wenn zur Bewertung des Kontextes **kein eigenes Wissen mit naturwissenschaftlichem Be-zug** hinzugezogen wird.	**N=0**
	NW-FB-UG: Ungeprüft / Unbelegt (UG)	Zur Bewertung werden **nicht überprüfte Aus-sagen** mit **natur-wissenschaftli-chem Inhalt** her-angezogen (Hö-rensagen).	"Allein wegen Silizium-mangel, allein schon So-larenergie. Da gibt es ja auch Risiken und Ne-benwirkungen, die nicht optimal sind." (S3.I2-SS23-A-B2_3-BN10LB, Pos. 9)	Erfüllt, wenn **Informationen aus unwissenschaftlichen Quellen**, wie öffentliche Medien (Zeitung, Fernsehen, Radio ohne wissen-schaftlichen Schwerpunkt) oder das soziale Umfeld, zur Bewertung des Kontextes herangezogen werden, diese aber nicht selbst wissenschaft-lich überprüft wurden. Auch erfüllt, wenn die Information zwar einen naturwissenschaftlichen Bezug hat, aber **keine Angabe zur Quelle ge-macht wird.**	**N=1**
	NW-FB-FU: Fundiert (FU)	Zur Bewertung werden **über-prüfte Aussagen** mit **naturwissen-schaftlichem In-halt** herangezo-gen.	"Als allererstes die In-formation, die ich habe von Fracking, gingen am Anfang von YouTube, dem Kanal Kurzgesagt aus, gab es ein deutsches und ein englischsprachiges Vi-deo." (S5.I1-WS23-D-M01-SJ21BO, Pos. 27)	Erfüllt, wenn **Informationen aus wissenschaftlichen oder wissen-schaftsjournalistischen Quellen** bei der Bewertung des Kontextes Einzug finden.	**N=2**

Tabelle A 5: Kodierleitfaden S3.I1&2 + FU – Aspekte naturwissenschaftlichen Wissens (NW-A)

S3.I1&2 - Naturwissenschaftliches Wissen (NW)					
Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug fachliche Aspekte? (NW-A)					
Kontexteinheit:	**Kodiereinheit:**	**Abstraktionsniveau:**	**Strukturdimension:**		**Auswertungseinheit:**
Das Transkript zum jeweiligen, personenbezogenen Interview.	Klare, bedeutungstragende Elemente im Interview, wie Satzteile oder Wortzusammenhänge.	Generalisierbare Aussagen zu Faktoren, Kriterien oder Formulierungen, welche aufzeigen, welche Informationen zur Bewertung als relevant erachtet werden. Mittleres Abstraktionsniveau.	Explizit genannte Faktoren, Kriterien oder Formulierungen sowie implizite Hinweise, welche aufzeigen, welche Art von Informationen bei der Bewertung des Kontextes als relevant erachtet wurden.		Alle Interviewtranskripte der Studie S3.
Kategorie:		**Definition:**	**Ankerbeispiel:**	**Kordierregel:**	**Kodierung:**
Aspekte naturwissenschaftlicher Bewertung (A)	**NW-A-KB: Keine naturwissenschaftliche Bewertung (KB)**	Es findet keine Bewertung mittels naturwissenschaftlicher Aspekte statt.		Erfüllt, wenn keine Bewertung von Quellen mittels naturwissenschaftlicher Aspekte vorgenommen wird. Immer dann kodiert, wenn keine andere Kategorien der Untersuchungsfrage kodiert werden konnten.	Nicht erfüllt: 0 Erfüllt: 1
	NW-A-EF: Effizienz (EF)	Die **Effizienz** wird als relevanter Faktor angesehen.	"Was es einfach für Zahlen gibt und für Werte, also wie effizient diese Energiegewinnung ist." (S3.I1-SS23-A-B3_5-SJ16DI, Pos. 42)	Erfüllt, wenn die **Effizienz** implizit oder explizit als relevant für die Bewertung angesehen wird.	Nicht erfüllt: 0 Erfüllt: 1
	NW-A-KA: Kostenabwägung (KA)	Die **Kosten** oder das Kosten-Nutzen-Verhältnis wird als relevanter Faktor angesehen.	"Die Zusammensetzung der Lithiumakkus und auch wie, ich sage mal, wie, wie preisgünstig die auch produziert werden können und was E-Fuels und Wasserstoff angeht, […]" (S3.I1-SS23-B-B3_5-WJ32FE, Pos. 42)	Erfüllt, wenn die **Kosten** oder das **Kosten-Nutzen-Verhältnis** implizit oder explizit als relevant für die Bewertung angesehen wird.	Nicht erfüllt: 0 Erfüllt: 1
	NW-A-NA: Nachhaltigkeit und Auswirkungen (NA)	Die **Auswirkungen** und/oder **Nachhaltigkeit** wird als relevanter Faktor angesehen.	"Aber so den Punkt der Nachhaltigkeit, würde ich auf jeden Fall sagen, wäre bei mir so an oberster Stelle und dann würde ich halt schauen okay, was ist nachhaltiger, was ist dafür sinnvoller? Und was sind jetzt die Negativaspekte, auch wenn das Ding jetzt nachhaltiger ist und man die gegeneinander abwägt?" (S3.I2-SS23-B-B3_5_XJ36JE, Pos. 21)	Erfüllt, wenn **Auswirkungen auf gesundheitlicher sozialer oder ökologischer Nachhaltigkeitsebene** implizit oder explizit als relevant für die Bewertung angesehen werden.	Nicht erfüllt: 0 Erfüllt: 1
	NW-A-FT: Funktionsweise (FT)	Die **Funktionsweise** wird als relevanter Faktor angesehen.	"Dann noch von irgendwie den E-Autoherstellern wie das Ganze, wie der Herstellungsprozess da vonstattengeht, was da alles berücksichtigt werden muss, wie das Ganze funktioniert. Sich einfach mit der Theorie befassen, die hinter so einem E-Auto steckt oder hinter so ein Brennstoffzellenauto steckt, damit man einfach mal weiß, nicht nur, dass ich das Auto fahren kann, sondern dass ich auch weiß, wie es funktioniert." (S3.I2-SS23-B-B3_5-XJ27BO, Pos. 23)	Erfüllt, wenn das **Verstehen der Funktionsweise** oder der zugrundeliegenden Technik implizit oder explizit als relevant für die Bewertung angesehen wird. Auch erfüllt, wenn **Definitionen** für ein Grundverständnis erfragt werden.	Nicht erfüllt: 0 Erfüllt: 1
	NW-A-UM: Umsetzbarkeit (UM)	Die **Umsetzbarkeit** wird als relevanter Faktor angesehen.	"[…] vielleicht ist jetzt das eine super umweltfreundlich, aber ist das überhaupt realistisch? Oder ist das vielleicht einfach nur so ein Wunschdenken? Und kann das überhaupt in der Realität funktionieren?" (S3.I2-SS23-A-B3_5_EF28SH, Pos. 23)	Erfüllt, wenn die **Umsetzbarkeit** implizit oder explizit als relevant für die Bewertung angesehen wird. Auch erfüllt, wenn **Herausforderungen oder Barrieren** für die Umsetzung diskutiert werden.	Nicht erfüllt: 0 Erfüllt: 1

	NW-A-RH: Rohstoffherkunft (RH)	Die **Herkunft von Rohstoffen** und Bestandteilen wird als relevanter Faktor angesehen.	"[...] wo kommen die Rohstoffe her, die da eingebaut werden? Wenn man jetzt irgendwie alternative Akkus hätte, die man irgendwie in Deutschland produzieren könnte, mit Rohstoffen, Ressourcen, die durchaus mehr zur Verfügung stehen würden als Lithium, dann hätte ich gesagt auf jeden Fall E-Auto." (S3.I2-SS23-B-B3_5-XJ27BO, Pos. 13)	Erfüllt, wenn die **Herkunft von Rohstoffen oder Bauteilen** implizit oder explizit als relevant für die Bewertung angesehen wird.	Nicht erfüllt: 0 Erfüllt: 1
	NW-A-NO: Notwendigkeit (NO)	Die **Notwendigkeit** oder der **Bedarf** wird als relevanter Faktor angesehen.	"[…]inwieweit welche Energien am intensivsten gebraucht werden, bezogen auf Energiepolitik und inwieweit erneuerbare Energien ausgebaut wurden und noch werden." (S2.I1-Wi22_23-A2-B2_3-XJ27BO, Pos. 30)	Erfüllt, wenn der **Bedarf oder die Notwendigkeit** implizit oder explizit als relevant für die Bewertung angesehen wird. Auch erfüllt, wenn **Alternativen** erfragt werden, da dann eine implizite Notwendigkeitsabfrage vorliegt.	Nicht erfüllt: 0 Erfüllt: 1

Tabelle A 6: Kodierleitfaden S3.I1&2 + FU – Verständnis Selbstreflexion (SR-V)

S3.I1&2 - Selbstreflexion (SR)					
Welches Verständnis haben Chemielehramtsstudierende von dem Begriff der *Selbstreflexion*? (SR-V)					
Definition: Selbstreflexion ist ein Prozess der introspektiven Überprüfung eigener Gedanken, Gefühle, Verhaltensweisen und Prägungen (Motivation, soziales Umfeld, eigene Erfahrungen uvm.) sowie darauf aufbauend ein Bewusstsein für externe und interne Einflussfaktoren. Sie umfasst die Wahrnehmung, Gewichtung und Bewertung ebenjener, um diese in Entscheidungsprozessen berücksichtigen sowie abwägen zu können und so eine objektivere Position zu erhalten. (Carver & Scheier, 2001; Facione, 1990; Grant et al., 2002; Silvia, 2020).					
Kontexteinheit:	**Kodiereinheit:**	**Abstraktionsniveau:**		**Strukturdimension:**	**Auswertungseinheit:**
Das Transkript zum jeweiligen, personenbezogenen Interview.	Klare, bedeutungstragende Elemente im Interview, wie Satzteile oder Wortzusammenhänge.	Generalisierbare Aussagen zu Faktoren, Kriterien oder Formulierungen, welche das individuelle Verständnis des Begriffs der *Selbstreflexion* darlegen. Mittleres Abstraktionsniveau.		Auf explizite Nachfrage genannte Faktoren, Kriterien oder Formulierungen, welche das individuelle Verständnis des Begriffs der *Selbstreflexion* darlegen.	Alle Interviewtranskripte der Studie S3.
Kategorie:		**Definition:**	**Ankerbeispiel:**	**Kordierregel:**	**Niveau:**
Verständnis Selbstreflexion (V)	**SR-V-ND: Nicht definierbar (ND)**	*Selbstreflexion* kann **nicht definiert** werden oder es liegen **diffuse Definitionsansätze** vor.	"Reflexion wäre zum Beispiel man hat sich ein Produkt genehmigt, was man sich normalerweise hätte nicht genehmigen wollen, aber da die Zeit nicht da war, hat man sich das halt gekauft, zum Beispiel." (S2.I1-Wi22_23-A2-B2_3-XJ27BO, Pos. 78)	Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diffuse oder unverständliche Definitionsansätze zu dem Begriff *Selbstreflexion* vorliegen.	**N=0**
	SR-V-IT: Implizites Teilverständnis (IT)	Der Begriff der *Selbstreflexion* liegt mit einer **naiven Definition** vor, welche aber **implizit** einen **Teilbereich der Grunddefinition** enthält.	AB: "Reflektiert heißt kritisch damit umgehen. Bedeutet, wenn ich, ich habe eine Meinung dazu und ich rede mit jemandem darüber. Meistens tauscht man sich ja aus und wenn jemand eine andere Ansicht hat, dann gehst du nicht erst mal hin und sagst "Nein, meine Meinung ist so, deine Meinung ist so, nein da kommen wir nicht zusammen. Das funktioniert nicht." (S2.I1-Wi22-B-M03-NÜ20FG, Pos. 58) SV: "Aber nicht nur noch mal generell drüber nachdenken, sondern auch immer zu gucken, wie kann ich nochmal vielleicht was verbessern oder was habe ich auch gut gemacht oder schlecht natürlich auch." (S2.I1-Wi22_23-B-B2_3-XF21FM, Pos. 98) EM: "Und, ja, bei der Reflexion allgemein, dass man vielleicht auch wie jetzt das Beispiel, was ich mit meinen Eltern genannt habe, auch da versucht, sich in andere Personen einfach hineinzuversetzen und halt auch aus deren Sicht praktisch versucht zu reflektieren, um einfach das besser verstehen zu können, warum die halt so denken, warum sie so denken." (S3.I2-SS23-B-B3_5_XJ36JE, Pos. 53) HF: "Ja einmal so kritisch hinterfragen, finde ich, passt da ganz gut." (S3.I1-SS23-B-B3_5-XJ19MF, Pos. 73)	**Subkategorien:** **AB: Abwägen** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als ein **abwägender Prozess zwischen zwei Standpunkten** verstanden wird. **SV: Selbstverbesserung** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als ein **Prozess der Selbstverbesserung** oder **Selbstkritik** verstanden wird. **EM: Empathie** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als das **Hineinfühlen in andere Personen** verstanden wird. Ebenfalls erfüllt, wenn die **Auswirkungen des eigenen Handelns auf andere Personen** hinterfragt wird. **HF: Hinterfragen** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als das **Hinterfragen von Themen oder Dingen** verstanden wird, ohne den Prozess weiter auszuführen.	**N=1**

Verständnis Selbstreflexion (V)	**SR-V-ET: Explizites Teilverständnis (ET)**	Der Begriff der *Selbstreflexion* liegt mit einer **informierten Definition** vor, welcher **explizit** einen **Teilbereich der Grunddefinition** enthält.	IS: "Und vielleicht auch was, viel fühlen, finde ich. Also wenn man reflektiert, so finde ich auch, wie habe ich mich da gefühlt? Habe ich mich wohlgefühlt? Habe ich mich unwohl gefühlt?. So auch zu reflektieren wie empfinde ich gerade? Wo stehe ich gerade so ein bisschen?" (S2.I1-Wi22_23-B-B2_3-BN10LB, Pos. 66) HA: "Ja eben, dass man halt noch mal sich vielleicht auch so widerspiegelt oder vielleicht auch mal über sein Handeln und seine Aussagen nachdenkt und vielleicht auch mal nicht immer sagt ja, aber meine Meinung ist die einzig wahre oder ich mache alles richtig, sondern auch mal denkt okay, das war jetzt nicht so gut, was du getan hast oder da hättest du vielleicht irgendwie verständnisvoller sein können oder ähnliches." (S3.I1-SS23-A-B3_5_IP17SE, Pos. 71) EX: "Im Endeffekt so auseinandernehmen oder nachverfolgen der Einflüsse und Gegebenheiten, die dahinter stehen können könnten, Also hinter den Informationen, hinter der Aufbereitung, hinter der Präsentation der Informationen und auch hinter meiner Wahrnehmung der Information. (...) Also wie ich vor-beeinflusst sein könnte, aber auch wie die Menschen, die die Information präsentieren, vor-beeinflusst sein könnten." (S3.I2-SS23-B-B3_5_SJ16DI, Pos. 21)	**<u>Subkategorien:</u>** **IS: Introspektion** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als der **Zugang zu seinen eigenen Emotionen, Gedanken oder Werten** oder die **Fähigkeit zur Introspektion** benannt wird. **HA: Überdenken von Handlungen** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als die **ex post Bewertung eigenen Handelns** oder Verhaltens verstanden wird. **EX: Bewusstsein für externe Prägungen** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als das **Hinterfragen von externen prägenden Faktoren** verstanden wird.	**N=2**

Tabelle A 7: Kodierleitfaden S3.I1&2 + FU – Aspekte der Selbstreflexion (SR-A)

S3.I1&2 - Selbstreflexion (SR)					
Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug selbstreflexive Aspekte? (SR-A)					
Definition: Selbstreflexion ist ein Prozess der introspektiven Überprüfung eigener Gedanken, Gefühle, Verhaltensweisen und Prägungen (Motivation, soziales Umfeld, eigene Erfahrungen uvm.) sowie darauf aufbauend ein Bewusstsein für externe und interne Einflussfaktoren. Sie umfasst die Wahrnehmung, Gewichtung und Bewertung ebenjener, um diese in Entscheidungsprozessen berücksichtigen sowie abwägen zu können und so eine objektivere Position zu erhalten. (Carver & Scheier, 2001; Facione, 1990; Grant et al., 2002; Silvia, 2020).					
Kontexteinheit:	**Kodiereinheit:**	**Abstraktionsniveau:**	**Strukturdimension:**		**Auswertungseinheit:**
Das Transkript zum jeweiligen, personenbezogenen Interview.	Klare, bedeutungstragende Elemente im Interview, wie Satzteile oder Wortzusammenhänge.	Generalisierbare Aussagen zu Faktoren, welche einen Einfluss auf das eigene Verhalten, Denken oder Bewerten haben könnten. Hohes Abstraktionsniveau.	Sowohl explizite Nennungen von als auch implizite Hinweise auf Faktoren, Kriterien oder die Formulierung von Fragen, welche bei der Bewertung von Problemstellungen mit naturwissenschaftlichem Bezug einen selbstreflexiven Anteil nahelegen.		Alle Interviewtranskripte der Studie S3.
Kategorie:		**Definition:**	**Ankerbeispiel:**	**Kordierregel:**	**Kodierung:**
Aspekte der Selbstreflexion (A)	**SR-A-KS: Keine Selbstreflexion (KS)**	Es sind **keine** oder nur **diffuse Anzeichen** von Selbstreflexion bei der Bewertung festzustellen.		Erfüllt, wenn **keine Selbstreflexion** vorliegt oder nur ein diffuses Gefühl genannt wird. Immer dann kodiert, wenn **keine andere Kategorie** dieser Untersuchungsfrage kodiert werden konnte.	Nicht erfüllt: 0 Erfüllt: 1
	SR-A-IF: Interne Einflussfaktoren (IF)	Es finden sich Hinweise auf eine **innere Einsichtnahme**, erkenntlich durch ein **Bewusstsein für selbstbezogene Faktoren,** wie Emotionen, Wissen u.Ä..	FW: "Aber halt auch so ein bisschen so, woher hatte ich das Vorwissen? Also, habe ich es aus der Schule? Habe ich es aus der Uni? Habe ich es aus einem Buch? Wenn ich das noch weiß." (S2.I1-Wi22_23-B-B2_3-BN10LB, Pos. 52) PR: "Also da müsste ich mich über unsere Möglichkeiten genauer informieren und dann eher die Vor- und Nachteile für mich persönlich abwägen, was vielleicht für mich persönlich wichtiger ist, aber auch für die Gesellschaft und die Umwelt." (S3.I1-SS23-B-B3_5-XJ19MF, Pos. 13) MW: "Ja, ich würde sagen, dass es einmal meine Normen und meine Werte, die ich habe, die beeinflussen ja auf jeden Fall meine Meinungsbildung, aber auch meine Einstellung natürlich zu unterschiedlichen Dingen, die das Thema dann beeinflussen, [...]" (S3.I1-SS23-B-B3_5-XJ19MF, Pos. 51)	**Subkategorien:** **FW: Fachwissen** Erfüllt, wenn ein Bewusstsein für die Prägung durch die **eigene fachliche Grundlage** ersichtlich wird. **PR: Persönliche Relevanz** Erfüllt, wenn eine Abwägung der **Vor- und Nachteile** für die eigene Person stattfindet oder die subjektive **selbstbezogene Notwendigkeit** angeführt wird. Ebenfalls erfüllt, wenn es sich um einen **persönlichen Möglichkeitsrahmen** handelt. **MW: Moral und Werte** Erfüllt, wenn ein Bewusstsein für den Einfluss der eigenen **Normen, Moralvorstellungen und Werten** ersichtlich wird. Ebenfalls erfüllt, wenn ein Bewusstsein für die Prägung durch **Selbstbild** oder der **Identifikation** stattfindet.	Nicht erfüllt: 0 Erfüllt: 1

Aspekte der Selbstreflexion (A)	**SR-A-IF: Interne Einfluss-faktoren (IF)**	Es finden sich Hinweise auf eine **innere Einsicht-nahme**, erkenntlich durch ein **Bewusst-sein für selbstbe-zogene Faktoren,** wie Emotionen, Wissen u.Ä..	HA: "Und dann lasse ich das auch noch mal im Kopf Revue passieren, was da, worüber man geredet hat. Überleg manchmal, je nachdem. Vor allem, wenn es jetzt irgendwie dazu kam, dass ich vielleicht einer Person Unrecht getan habe "Ja, kam es dazu? Was hat diese Situation, ja, hervorgerufen? Was war da vielleicht mein Fehlverhalten? Wie kann ich das besser machen?" Und es läuft auch meistens darauf hinaus, dass ich dann noch mal mit den Menschen, die betroffen sind, darüber rede und versuche das zu klären. Bevor da jetzt irgendwelche Probleme entstehen." (S3.I1-SS23-B-B3_5-EF28SH, Pos. 59) EM: "[...] also so argumentiere, rede und auch zu dem Sachverhalt selbst, ob ich dann emotional irgendwie da betroffen bin." (S3.I2-SS23-A-B2_3-BN10LB, Pos. 37) EF: "Ja und dann halt einfach so wie ich das mich persönlich, also ob ich persönliche Erfahrungen damit habe, [...]" (S3.I2-SS23-A-B2_3-BN10LB, Pos. 13) IN: "Also dann wäre es mir schon wichtig, dass sich das nicht an solchen Sachen festmache wie persönlichem Interesse oder so Neigungen irgendwie." (S3.I1-SS23-A-B3_5-SJ16DI, Pos. 54) BV: "Ja erstmal, ob ich persönlich befangen bin. So ein bisschen. Weil es gibt natürlich auch Themen, die für mich persönlich bedeutsam sind und ob das meine Wahl dann beeinflusst. Also inwiefern, meine persönliche Meinung da eine Rolle spielt." (S2.I1-Wi22_23-B-B2_3-BN10LB, Pos. 52)	**HA: Handlungen** Erfüllt, wenn eine **ex post Betrachtung eigenen Handelns oder Standpunkts** stattfindet oder diese als relevanter Vorgang angesehen wird.. **EM:** E**motionen** Erfüllt, wenn ein Einblick in die **eigene Gefühlswelt** stattfindet oder die Prägung durch Emotionen thematisiert wird. **EF: Erfahrungen** Erfüllt, wenn **eigene Erfahrungen** als Grundlage zur Bewertung herangezogen werden oder ein Bewusstsein für deren prägenden Charakter ersichtlich wird. Ebenfalls erfüllt, wenn die **Gewöhnung** angeführt wird, da diese aus (häufig) gemachten Erfahrungen hervorgeht. **IN: Interesse** Erfüllt, wenn ein Einblick in das eigene **(Des-)Interesse** bei der Bewertung herangezogen wird oder ein Bewusstsein für dessen Einfluss ersichtlich wird. **BV: Befangenheit, Vorurteile** Erfüllt, wenn ein Bewusstsein für den Einfluss von **Vorurteilen oder Vormeinungen** ersichtlich wird. Darunter fällt auch **Sympathie und das Versteifen auf Sichtweisen oder Meinungen oder Confirmation Bias.**	Nicht erfüllt: 0 Erfüllt: 1

Aspekte der Selbstreflexion (A)	**SR-A-IF: Interne Einfluss-faktoren (IF)**	Es finden sich Hinweise auf eine **innere Einsicht-nahme**, erkenntlich durch ein **Bewusst-sein für selbstbe-zogene Faktoren,** wie Emoti-onen, Wis-sen u.Ä..	PK: "Und ja ich halt als Person auch ich bin und nicht wer anders. Das ist, denke ich glaub ich, auch ein großer Punkt. Also ich gehe davon aus, dass jede Person irgendwie in irgendeiner Form eine Besonderheit an sich hat, die sie zu der Person macht, die sie ist, die sie oder er ist. Und daher denke ich, dass jede Person auch gewisse Dinge mitbringt, die jetzt die die anderen Menschen zum Beispiel nicht haben. Sei es irgendein Meinungsbild, irgendein Erlebnis, irgendein Trauma, was die Person vielleicht erlebt hat. Das kann ja auch beeinflussen, wie man sich eine Meinung bildet." (S5.I2-WS23-D-M01-MB14BO, Pos. 46) KV: " [...] dass Menschen subjektiv wahrnehmen und gebiased sind und wie viele Biases es davon gibt." (S3.I2-SS23-B-B3_5_RI05CH, Pos. 37)	**PK: Persönlichkeit** Erfüllt, wenn die **Persönlich-keit oder eigene Identität** als beeinflussend wahrgenom-men oder reflektiert wird. **KV: Kognitive Verzerrun-gen** Erfüllt, wenn **kognitive Ver-zerrungen** als beeinflussend wahrgenommen oder reflek-tiert werden.	Nicht erfüllt: 0 Erfüllt: 1
Aspekte der Selbstreflexion (A)	**SR-A-EF: Externe Einfluss-faktoren (EF)**	Es finden sich Hin-weise auf ein Be-wusstsein für **beein-flussende Faktoren** beim Tref-fen von Entschei-dungen. Dieses Be-wusstsein muss **expli-zit** deutlich gemacht werden.	SN: "Genauso würde ich aber auch im sozialen Umfeld sagen klar, wir lassen uns irgendwie immer unbewusst beeinflussen, auch wenn man es vielleicht gar nicht möchte." (S2.I1-Wi22_23-B-B2_3-XF21FM, Pos. 86) ME: "Nein, abgesehen von meinem Vorwissen, sondern die Medien haben sicherlich viel Einfluss, wie darüber berichtet wird, wobei ich auch der Meinung bin, dass es durchaus, Äh ja, dass das Medien tendenziös tendenziös manchmal berichten. Nicht immer, aber manchmal, so zumindest mein, mein Gefühl, sage ich mal, oder wie ich manchmal an der Berichterstattung grad zu E-Mobilität sehe oder auch zu Wasserstoff. Aber Medien haben da sicherlich einen Einfluss [...]" (S3.I1-SS23-B-B3_5-WJ32FE, Pos. 54) BI: "Jetzt, ich glaube, die Uni würde mich schon stark beeinflussen, weil es für mich die höchste Autorität einfach ist. Das ist für mich eigentlich wissenschaftlich. Und deswegen glaube ich dem auch (lacht)." (S3.I1-SS23-A-B3_5-SV32BT, Pos. 63) PO: "Also ich glaube aus, aus so einer, wie jetzt häufig ja so gerne so so diese linksgrünversiffte mit ganz, ganz großen Anführungszeichen, aber das ist ja so das Wort, was man da manchmal hört. So, so Meinungen aus oder Ansichten aus der Richtung, die prägen mich mittlerweile, glaube ich, relativ stark." (S3.I2-SS23-A-B3_5_EF28SH, Pos. 23) RE: "Ich glaube, dass so was wie Glaube oder so einen häufig beeinflusst in sowas." (S5.I1-WS23-C-M01-VM14OL, Pos. 58)	**<u>Subkategorien:</u>** **SN: Soziale Norm** Erfüllt, wenn ein Bewusstsein für Prägungen seitens des **so-zialen Umfeldes** ersichtlich wird. Dies schließt **soziale Erwünschtheit** und **soziale Kontexte,** wie das Arbeitsum-feld, mit ein, ebenso wie Ein-flüsse von **Freunden oder Familie**. Darunter fallen ebenfalls Einflüsse aus der **Kindheit** oder Prägungen durch die **Erziehung,** ebenso wie **Prägungen durch Her-kunft.** **ME: Medien** Erfüllt, wenn ein Bewusstsein für Prägungen durch **Medien** ersichtlich wird. **BI: Bildung** Erfüllt, wenn ein Bewusstsein für die Prägung durch **Bil-dung und Ausbildung** er-sichtlich wird. Dies schließt die Prägung durch **Autoritä-ten** und **Expertenmeinun-gen** mit ein, welche im Bil-dungskontext entstehen. **PO: Politik** Erfüllt, wenn ein Bewusstsein für Prägungen durch **politi-sche Vertreter oder Politik** ersichtlich wird. **RE: Religion** Erfüllt, wenn **Religion** als be-einflussend wahrgenommen oder reflektiert wird.	Nicht erfüllt: 0 Erfüllt: 1

Tabelle A 8: Kodierleitfaden S3.I1&2 + FU – Aspekte der Epistemologie (EP-A)

S3.I1&2 - Epistemologisches Verständnis (EP)				
Inwieweit finden sich bei Chemielehramtsstudierenden bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug epistemologische Aspekte? (EP-A)				
Epistemologie ist die Untersuchung der Natur, des Ursprungs und des Umfangs des Wissens. Sie befasst sich mit den Bedingungen, unter denen Überzeugungen als gerechtfertigt betrachtet werden und wie Wahrheitsansprüche geprüft werden können. In der wissenschaftlichen Praxis umfasst die Epistemologie die Kriterien und Methoden, die zur Bewertung der Glaubwürdigkeit, Zuverlässigkeit und Validität von wissenschaftlichen Erkenntnisse und Theorien herangezogen werden.				
Kontext-einheit:	**Kodiereinheit:**	**Abstraktionsniveau:**	**Strukturdimension:**	**Auswertungs-einheit:**
Das Transkript zum jeweiligen, personenbezogenen Interview.	Klare, bedeutungstragende Elemente im Interview, wie Satzteile oder Wortzusammenhänge.	Generalisierbare Aussagen zu Faktoren, welche einen Einfluss auf das eigene Verhalten, Denken oder Bewerten haben könnten. Hohes Abstraktionsniveau.	Sowohl explizite Nennungen von als auch implizite Hinweise auf Faktoren, Kriterien oder die Formulierung von Fragen, welche bei der Bewertung von Problemstellungen mit naturwissenschaftlichem Bezug einen selbstreflexiven Anteil nahelegen.	Alle Interviewtranskripte der Studie S3.

Kategorie:		**Definition:**	**Ankerbeispiel:**	**Kordierregel:**	**Kodierung:**
(A) Aspekte des epistemologischen Verständnisses	**EP-A-KB: Keine Bewertung mittels epistemologischer Aspekte (KB)**	Es findet **keine Bewertung** mittels epistemologischer Aspekte statt.		Erfüllt, wenn **keine Bewertung von Quellen** mittels epistemologischer Aspekte vorgenommen wird. Immer dann kodiert, wenn **keine andere Kategorie** der Untersuchungsfrage kodiert werden konnten.	Nicht erfüllt: 0 Erfüllt: 1
	EP-A-B: Bewertung mittels epistemologischer Aspekte (B)	Bei der Bewertung finden **epistemologische Aspekte Verwendung**, um bspw. Quellen oder die Genese von Informationen einschätzen zu können.	EX: "Also die Meinung der, wirklich der, der der Wissenschaftler, die sich damit hauptberuflich auseinandersetzen, und da alle Pro- und Contra Argumente vielleicht schon mal formuliert haben." (S3.I2-SS23-A-B3_5_VI21ED, Pos. 19) AK: "Wann wurde das gemacht? Also die Zeit, in welcher Zeit das gemacht wurde spielt ja auch eine Rolle." (S3.I2-SS23-A-B2_3-BN10LB, Pos. 51) IK: "Es gibt ja liberale, konservative Zeitungen, die dann dementsprechend auch Meinungen so oder so vertreten, oder auch bei gewissen Autoren ist es bekannt, dass die absolut diese Stellung immer einnehmen [...]" (S3.I2-SS23-A-B3_5_WJ32FE, Pos. 25)	**Subkategorien:** **EX: Expertise der Autor*in** Erfüllt, wenn die **Autorität und Expertise der*s Autor*in** als relevante Faktoren bei der Einschätzung der Vertrauenswürdigkeit einer Quelle genannt werden. **AK: Aktualität** Erfüllt, wenn die **Aktualität der Veröffentlichung** als relevanter Faktor bei der Einschätzung der Vertrauenswürdigkeit einer Quelle genannt wird. **IK: Interessenskonflikte** Erfüllt, wenn etwaige **Interessenskonflikte** oder **Subjektivität (Meinungen)** als relevanter Faktor bei der Einschätzung der Vertrauenswürdigkeit einer Quelle genannt werden. Auch die umgekehrte Formulierung der gewünschten Unabhängigkeit von Quellen oder Berichterstattern werden hier kodiert.	Nicht erfüllt: 0 Erfüllt: 1

(A) Aspekte des epistemologischen Verständnisses	**EP-A-B: Bewertung mittels epistemologischer Aspekte (B)** **EP-A-B: Bewertung mittels epistemologischer Aspekte (B)**	Bei der Bewertung finden **epistemologische Aspekte Verwendung**, um bspw. Quellen oder die Genese von Informationen einschätzen zu können.	PR: "Ja, also als allererstes erstmal der Ton, in dem da geschrieben wird. Also muss jetzt, ich spreche jetzt gerade noch nicht von irgendwelchen wissenschaftlichen Studien usw. und so fort, sondern erst mal nur von dem, was man meistens bekommt, wenn man halt bei einer Suchmaschine eingibt, so und so, halt irgendwelche journalistischen Reporte da drauf, da achte ich darauf, dass die der Ton, in dem das geschrieben ist, möglichst neutral, möglichst sachlich, und dass da irgendwie nicht direkt eine Positionierung für die eine oder andere Richtung mitschwingt." (S3.I2-SS23-A-B3_5_EF28SH, Pos. 15)	**PR: Präsentation** Erfüllt, wenn die **Präsentation** oder **Sprache und Stil der Quelle** als relevante Faktoren bei der Einschätzung der Vertrauenswürdigkeit einer Quelle genannt werden. Auch erfüllt, wenn **Schlüssigkeit und Logik** der Argumente angeführt werden.	Nicht erfüllt: 0 Erfüllt: 1
			HK: "Das Ding ist so, die die großen Quellen, eben halt Nachrichten in den ARD und ZDF oder auch die Zeitungsbeiträge, Artikel, die einem grundsätzlich erstmal seriös erscheinen, von großen Zeitungen, von also, oder Seiten, die nur über Nachrichten berichten, die erscheinen oder den glaub ich grundsätzlich erstmal schon, weil ich denke, dass die seriös sind und dass sie in Deutschland, man mit der Pressefreiheit so schon, das die den größten Teil halt korrekt berichten." (S3.I2-SS23-B-B3_5_IP17SE, Pos. 41)	**HK: Herkunft** Erfüllt, wenn die **Herkunft** oder **Veröffentlichungsplattform** sowie deren Bekanntheit als relevante Faktoren bei der Einschätzung der Vertrauenswürdigkeit einer Quelle genannt werden.	
			ME: "Und dann von da aus weitergehend halt zu schauen "okay, jetzt haben wir diese Quellen. Gucken wir da mal weiter. Was ist denn das? Was ist denn das für eine Studie? Wie wurde die durchgeführt? Was sind die Methoden?" (S3.I2-SS23-A-B3_5_EF28SH-Post, Pos. 15)	**ME: Methodik** Erfüllt, wenn die **Methodik und** das **Forschungsdesign** als relevanter Faktor bei der Einschätzung der Vertrauenswürdigkeit einer Quelle genannt werden	
			KO: "Aber ich glaube, ich würde schon darauf achten, verschiedene Quellen hin, also bei der Recherche hinzuzuziehen. Und nicht nur die Windkraftexperten oder nicht nur die Solarenergieexperten, oder Befürworter, also die Quellen zu nutzen, sondern ich glaube, ich würde schon darauf achten, dass es so breit gefächert ist." (S3.I1-SS23-A-B3_5-SJ16DI, Pos. 44)	**KO: Konsistenzprüfung** Erfüllt, wenn die **Übereinstimmung mit anderen Quellen** als relevanter Faktor bei der Einschätzung der Vertrauenswürdigkeit einer Quelle genannt werden. Ebenfalls erfüllt, wenn **andere Standpunkte bei der Bewertung** hinzugezogen werden, um eine bessere Aussagekraft zu erreichen hinzugezogen werden.	

(A) Aspekte des epistemologischen Verständnisses	**EP-A-B: Bewertung mittels epistemologischer Aspekte (B)** **EP-A-B: Bewertung mittels epistemologischer Aspekte (B)**	Bei der Bewertung finden **epistemologische Aspekte Verwendung**, um bspw. Quellen oder die Genese von Informationen einschätzen zu können.	TR: "Also genau, ich würde natürlich auch gucken, wie die Informationen aufbereitet sind, also wie, ob Studien mit Namen angegeben sind oder vielleicht verlinkt sind oder wie genau die Quellen angegeben sind und welche Quellen angegeben sind, ob das jetzt wissenschaftliche Publikationen sind oder andere Websites oder so was." (S3.I2-SS23-B-B3_5_SJ16DI, Pos. 17) SE: "Also man müsste dann auch gucken, wo wurde die Studie durchgeführt? Wenn er zum Beispiel jetzt auch ein Land ist, was jetzt sehr groß in Fracking ist, denke ich, wird das auch noch mal Einfluss darauf haben, was vielleicht auch in den Studien dann kommuniziert werden darf. Und halt auch wie, dass die Regierung generell so ist mit Zensur und sowas. Wenn ich nur an China denke, da werden auch häufig Studien oder sowas zitiert, zensiert, nicht zitiert, zensiert." (S5.I2-WS23-D-M01-MB14BO, Pos. 38)	**TR: Transparenz** Erfüllt, wenn deutlich wird, dass eine **offene und transparente Angabe von (wissenschaftlichen) (Erst-) Quellen oder Herkunftsinformationen** als relevante Faktoren bei der Einschätzung der Vertrauenswürdigkeit einer Quelle genannt werden. **SE: Soziokulturelle Eingebundenheit** Erfüllt, wenn bei der Bewertung einer Quelle die **soziokulturelle Eingebundenheit** der Erkenntnisse, des Autors oder der Plattform berücksichtigt werden.	Nicht erfüllt: 0 Erfüllt: 1

Tabelle A 9: Kodierleitfaden S3.I1&2 + FU – Verständnis Kritisches Denken (KD-V)

S3.I1&2 - Kritisches Denken (KD)					
Welches Verständnis haben Chemielehramtsstudierende von der Formulierung *kritisch denken* oder *kritisch sein*? (KD-V)					
Definition: Kritisches Denken (KD) bezeichnet eine sorgfältige und systematische sowie vor allem begründete Entscheidungsfindung bei der Auseinandersetzung der*s Denkenden mit einem Objekt oder Subjekt. Diese Auseinandersetzung beinhaltet die Evaluation und Anwendung des eigenen Wissens und der eigenen Fähigkeiten auf kognitiver Ebene sowie die Berücksichtigung der eigenen Emotionen, Werte, Haltung und Motivation auf affektiver Ebene. Es ermöglicht zudem eine konstante Selbstregulation und Anpassung, die eine Schleife zur*m Denkenden selbst zurückzieht und sie/ihn nicht nur in die Position der*s Bewertenden, sondern auch in die Position der*s zu Bewertenden versetzt (u. a. nach Rafolt, 2019; Paul & Elder, 2014, Facione, 1990).					
Kontext-einheit:	**Kodiereinheit:**	**Abstraktionsniveau:**	**Strukturdimension:**		**Auswertungs-einheit:**
Das Transkript zum jeweiligen, personenbezogenen Interview.	Klare, bedeutungstragende Elemente im Interview, wie Satzteile oder Wortzusammenhänge.	Generalisierbare Aussagen zu Faktoren, Kriterien oder Formulierungen, welche das individuelle Verständnis der Phrase *kritisch denken* oder *kritisch sein* darlegen. Mittleres Abstraktionsniveau.	Auf explizite Nachfrage genannte Faktoren, Kriterien oder Formulierungen, welche das individuelle Verständnis der Phrase *kritisch denken* oder *kritisch sein* darlegen.		Alle Interviewtranskripte der Studie S3.
Kategorie:		**Definition:**	**Ankerbeispiel:**	**Kordierregel:**	**Niveau:**
(V) Verständnis von Kritischem Denken	**KD-V-NA: Naive Vorstellung (NA)**	Die Formulierung *kritisch denken* oder *kritisch sein* ist **naiv** oder diffus definiert.	ND: "Ja, wenn man, wenn ich jetzt zu allem so eine, so eine starke Meinung hätte, aber jetzt nicht unbedingt im negativen Sinne, sondern ist ja schon eher neutraler, weil, also, man kann ja kritisch zu Sachen sein, die man ja auch gut findet, aber so, ja dem Begriff vom kritisch ist, ja finde ich schwierig." (S2.I1-Wi22_23-A1-B2_3-XJ36JE, Pos. 66) HF: "Eins das größte Überthema, was mir immer direkt ins Kopf liegt, ist hinterfragen." (S3.I2-SS23-B-B2_3-SV32BT, Pos. 49) SK: "Ja, vielleicht nicht immer alles direkt zu glauben. Also dass man auch vor allem, wenn irgendwas in den in den Nachrichten steht oder so." (S3.I1-SS23-A-B3_5_IP17SE, Pos. 67) SV: "Also ich übe sehr viel Selbstkritik so, das hättest du besser machen können oder das und das". (S2.I1-Wi22_23-B-B2_3-BN10LB, Pos. 60) TK: "Aber andere Menschen kritisiere ich nicht gerne oder mache es nicht häufig. Ich mache es nur, wenn ich es für angebracht oder sinnvoll erhalte. Also denke so, dass es Sinn bringt, die Person (unv.) zu kritisieren. Ich weiß nicht. Ich nehme meistens eher die Menschen so wie sie sind. Also das	**Subkategorien:** **ND: Nicht definierbar:** Erfüllt, wenn eine **diffuse oder keine Definition** von der Formulierung *kritisch denken* oder *kritisch sein* vorliegt. **HF: Hinterfragen:** Erfüllt, wenn bei Nachfrage nach dem Verständnis von der Formulierung *kritisch denken* oder *kritisch sein* dieses als **Hinterfragen von Dingen und Themen** verstanden wird, wobei dies **nicht weiter ausgeführt** wird. **SK: Skepsis** Erfüllt, wenn bei der Nachfrage nach dem Verständnis von der Formulierung *kritisch denken* oder *kritisch sein* dieses als die **Neigung** definiert wird, **nicht alles direkt zu glauben** oder **grundlegend skeptisch zu sein**. **SV: Selbstverbesserung/Kritik:** Erfüllt, wenn bei Nachfrage nach dem Verständnis von der Formulierung *kritisch denken* oder *kritisch sein* dieses als ein **Prozess der Selbstverbesserung** verstanden wird, welcher aus tadelnder Selbstkritik hervorgeht.	**N=0**

(V) Verständnis von Kritischem Denken	**KD-V-NA: Naive Vorstellung (NA)**	Die Formulierung *kritisch denken* oder *kritisch sein* ist **naiv** definiert.	sind wirklich kritische Sachen, wie jemand übt Rassismus aus oder Ungleichheiten, dann sag ich da schon was." (S2.I1-Wi22_23-B-B2_3-BN10LB, Pos. 60) SD: "Ja, das Erste, was mir da immer in den Kopf kommt, ist so aufklärerische Gedanken, also halt selbstständiges Denken und, aber halt im erweiterten Sinne, weil ich da ja jetzt. Ja, das ist so ein, wahrscheinlich auch wieder durch die Chemie sehr geprägter Begriff, also rational und faktisch orientiert denken, handeln." (S3.I1-SS23-B-B3_5-EF28SH, Pos. 49)	**TK: Tadeln bzw. Kritisieren:** Erfüllt, wenn bei der Nachfrage nach dem Verständnis von der Formulierung *kritisch denken* oder *kritisch sein* dieses als **kritisieren** im Sinne von **tadeln** anderer definiert wird. **SD: Selbstständiges Denken** Erfüllt, wenn bei der Nachfrage nach dem Verständnis von der Formulierung *kritisch denken* oder *kritisch sein* dieses als **eigenständiges Denken** definiert wird oder die eigene Beschäftigung mit Themen angeführt wird.	**N=0**
	KD-V-TG: Teilbereich der Grunddefinition (TG)	Unter der Formulierung *kritisch denken* oder *kritisch sein* wird ein **Teilbereich der Grunddefinition** verstanden.	SR: "Dieser sehr aktive Denkprozess. Das aktive sich mit eigenen Ansichten auseinanderzusetzen und vor allem halt auch zu hinterfragen." (S3.I2-SS23-A-B3_5_EF28SH, Pos. 27) WI: "Also Kritisches Denken ist für mich nicht nur irgendwie so Faktenwissen abrufen können, weil das kann das Internet ja auch, sondern dass man sich ein gewisses Wissen angeeignet hat, mit dem man jetzt irgendwie zu einer begründeten Stellungnahme kommt, bei der man Pro und Contra mit reinbringen kann. Dass man jetzt sagen kann: Also ungeachtet dessen, wie ich dazu jetzt stehe. Dass ich sage: okay, E-Autos, das sind Pro, das ist Contra, dass man beide Argumente gleichwertig hinstellt [...]" (S3.I2-SS23-B-B3_5-XJ27BO, Pos. 47) EP: "Auch wenn ich irgendwelche Statistiken sehe, dass ich halt gucke, woher kommt das? Wer hat das gemacht? Wann wurde das gemacht? Also die Zeit, in welcher Zeit das gemacht wurde spielt ja auch eine Rolle. Der Ort halt irgendwo auch, ob es jetzt USA ist, Deutschland ist auch nochmal ein Unterschied von der Gesellschaft." (S3.I2-SS23-A-B2_3-BN10LB, Pos. 51)	**<u>Subkategorien:</u>** **SR: Selbstreflexion** Erfüllt, wenn bei der Nachfrage nach dem Verständnis von der Formulierung *kritisch denken* oder *kritisch sein* dieses als das **Hinterfragen der eigenen Person oder den eigenen Prägungen** im Sinne einer **Selbstreflexion** verstanden wird. **WI: Wissen** Erfüllt, wenn bei der Nachfrage nach dem Verständnis von der Formulierung *kritisch denken* oder *kritisch sein* dieses als das **Hinterfragen und Hinzuziehen von externen Informationen oder eigenem Wissen** verstanden wird. Ebenfalls erfüllt, wenn deutlich wird, dass **faktenorientiertes Denken und Wissen** zum Kritischen Denken gehört. **EP: Epistemologisches Verständnis** Erfüllt, wenn bei der Nachfrage nach dem Verständnis von der Formulierung *kritisch denken* oder *kritisch sein* dieses als das **Hinterfragen und Bewerten von Quellen** und ihrer Glaubwürdigkeit verstanden wird.	**N=1**
	KD-V-MT: Zusammenspiel mehrerer Teilbereiche der Grunddefinition (MT)	Die Formulierung *kritisch denken* oder *kritisch sein* wird als Zusammenspiel von Selbstreflexion, epistemologischem Verständnis bzw. Fachwissen verstanden.		Erfüllt, wenn bei der Nachfrage nach dem Verständnis von der Formulierung *kritisch denken* oder *kritisch sein* dieses als **Zusammenspiel von mindestens zwei der Teilbereiche der Grunddefinition** genannt wird: (Fach-)Wissen, epistemologisches Verständnis (Quellenbewertung) und Selbstreflexion. Immer dann kodiert, wenn mindestens zwei der Subkategorien von TG kodiert wurden.	**N=2**

Tabelle A 10: Kodierleitfaden S3.I1&2 + FU – Fähigkeit zur Bewertung (FB)

S3.I1&2 - Fähigkeit zur Bewertung (FB)					
Inwieweit sehen sich die Studierenden des Chemielehramts in der Lage, eine klare Position bei der Bewertung des Kontextes einzunehmen?					
Kontext-einheit:	**Kodiereinheit:**	**Abstraktionsniveau:**	**Strukturdimension:**		**Auswertungs-einheit:**
Das Transkript zum jeweiligen personenbezogenen Interview.	Klare, bedeutungstragende Elemente im Interview, wie Satzteile oder Wortzusammenhänge.	Generalisierbare Aussagen zu Faktoren, Kriterien oder Formulierungen, welche aufzeigen, wie das eigene Fachwissen subjektiv eingeschätzt wird. Geringes Abstraktionsniveau.	Auf explizite Nachfrage genannte Faktoren, Kriterien oder Formulierungen, welche aufzeigen, ob eine subjektiv empfundene Fähigkeit zur Bewertung vorliegt.		Alle Interviewtranskripte der Studie S3.
	Kategorie:	**Definition:**	**Ankerbeispiel:**	**Kordierregel:**	**Niveau:**
(FB) Fähigkeit zur Bewertung	**FB-NL: Sieht sich nicht in der Lage (NL)**	Es wird sich **nicht zugetraut**, eine Entscheidung treffen zu können.	"Auf jeden Fall nicht, Nein. Gerade vor allem einerseits, weil es halt auch lange her ist. Und wenn es jetzt sage ich mal auch eine größere Entscheidung ist, wie ja zum Beispiel das, was ja eine sehr große Entscheidung ist, dann könnte ich so was auf keinen Fall auf die Schnelle beantworten." (S2.I1-Wi22_23-A1-B2_3-XJ36JE, Pos. 42)	Erfüllt, wenn sich auf Nachfrage **nicht zugetraut wird**, eine Entscheidung treffen zu können.	**N=0**
	FB-UE: Sieht sich unter Einschränkungen in der Lage (UE)	Es wird sich **mit Einschränkungen zugetraut**, eine Entscheidung treffen zu können. Entscheidungstendenzen liegen vor.	"Ich glaube, eine 100-prozentig fundamentierte Meinung kann ich nicht sagen, aber ich glaube. Also, ich könnte eine erste Tendenz geben, aber ich würde nicht davon sprechen, dass ich jetzt eine 100-prozentige Meinung dazu abgebe." (S5.I1-WS23-D-M01-VF42IB, Pos. 34)	Erfüllt, wenn sich auf Nachfrage nur **unter Einschränkungen zugetraut** wird, eine Entscheidung treffen zu können oder Entscheidungstendenzen vorliegen.	**N=1**
	FB-IL: Sieht sich in der Lage (IL)	Es wird sich **zugetraut**, eine Entscheidung treffen zu können.	"Ich meine, dem Wissen sind nie Grenzen gesetzt, sag ich mal, auch in diesem Thema nicht. Ich würde mich jetzt auch nicht als Fachmann bezeichnen, das muss ich ehrlicherweise auch sagen, aber ich würde schon sagen, dass ich für ja, ja, na, ich, doch ich würde schon sagen, dass ich für eine abschließende Meinung da schon ein gutes Wissensfundus habe." (S3.I1-SS23-B-B3_5-WJ32FE, Pos. 44)	Erfüllt, wenn sich auf Nachfrage **zugetraut** wird, eine Entscheidung treffen zu können.	**N=2**

A.3.1.2 Kodierleitfäden S3.F1&2

Tabelle A 11: Kodierleitfaden S3.F1&2 – Verständnis Selbstreflexion (VS)

S3.F1&2 - Verständnis Selbstreflexion (VS)

VS: Welches Verständnis haben Studierende des Chemielehramtes von dem Begriff *Selbstreflexion*?

Definition:
Selbstreflexion ist ein Prozess der introspektiven Überprüfung eigener Gedanken, Gefühle, Verhaltensweisen und Prägungen (Motivation, soziales Umfeld, eigene Erfahrungen uvm.) sowie darauf aufbauend ein Bewusstsein für externe und interne Einflussfaktoren. Es umfasst die Wahrnehmung, Gewichtung und Bewertung ebenjener, um diese in Entscheidungsprozessen berücksichtigen sowie abwägen zu können und so eine objektivere Position zu erhalten (Carver & Scheier, 2001; Facione, 1990; Grant et al., 2002; Silvia, 2020).

Kontexteinheit:	**Kodiereinheit:**	**Abstraktionsniveau:**	**Strukturdimension:**	**Auswertungseinheit:**
Die Antwort auf die zugehörige offene Fragestellung von VF1.	Klare, bedeutungstragende Elemente in der Antwort, wie Satzteile oder Wortzusammenhänge.	Generalisierbare Aussagen, welche das individuelle Verständnis des Begriffs der *Selbstreflexion* deutlich machen. Mittleres Abstraktionsniveau.	Sowohl explizite Nennungen als auch implizite Hinweise, die das individuelle Verständnis des Begriffs der *Selbstreflexion* verdeutlichen.	Alle Antworten auf die offene Fragestellung VF1 in den zugehörigen Fragebögen zu Studie 3.

Kategorie:		**Definition:**	**Ankerbeispiel:**	**Kodierregel:**	**Niveau:**
Verständnis Selbstreflexion (VS)	**VS-ND: Nicht definierbar (ND)**	*Selbstreflexion* kann **nicht definiert** werden oder es liegen **diffuse Definitionsansätze** vor.	"Reflexion wäre zum Beispiel man hat sich ein Produkt genehmigt, was man sich normalerweise hätte nicht genehmigen wollen, aber da die Zeit nicht da war, hat man sich das halt gekauft, zum Beispiel." (EH1-Wi22_23-A2-B2_3-XJ27BO-Pre, Pos. 78)	Erfüllt, wenn der Begriff der *Selbstreflexion* **diffus oder unverständlich** definiert wird.	**N=0**
	VS-IT: Implizites Teilverständnis (IT)	Der Begriff der *Selbstreflexion* liegt mit einer **naiven Definition** vor, welche aber implizit einen **Teilbereich der Grunddefinition** enthält.	AB: "Reflektiert heißt kritisch damit umgehen. Bedeutet, wenn ich, ich habe eine Meinung dazu und ich rede mit jemandem darüber. Meistens tauscht man sich ja aus und wenn jemand eine andere Ansicht hat, dann gehst du nicht erst mal hin und sagst "Nein, meine Meinung ist so, deine Meinung ist so, nein da kommen wir nicht zusammen. Das funktioniert nicht." (S2.I1-Wi22-B-M03-NÜ20FG, Pos. 58) SV: "Aber nicht nur noch mal generell drüber nachdenken, sondern auch immer zu gucken, wie kann ich nochmal vielleicht was verbessern oder was habe ich auch gut gemacht oder schlecht natürlich auch." (S2.I1-Wi22_23-B-B2_3-XF21FM, Pos. 98) EM: "Und, ja, bei der Reflexion allgemein, dass man vielleicht auch wie jetzt das Beispiel, was ich mit meinen Eltern genannt habe, auch da versucht, sich in andere Personen einfach hineinzuversetzen und halt auch aus deren Sicht praktisch versucht zu reflektieren, um einfach das besser verstehen zu können, warum die halt so denken, warum sie so denken." (S3.I2-SS23-B-B3_5_XJ36JE, Pos. 53) HF: "Ja einmal so kritisch hinterfragen, finde ich, passt da ganz gut." (S3.I1-SS23-B-B3_5-XJ19MF, Pos. 73)	**Subkategorien:** **AB: Abwägen** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als ein **abwägender Prozess zwischen zwei Standpunkten** verstanden wird. **SV: Selbstverbesserung** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als ein **Prozess der Selbstverbesserung** oder **Selbstkritik** verstanden wird. **EM: Empathie** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als das **Hineinfühlen in andere Personen** verstanden wird. Ebenfalls erfüllt, wenn die **Auswirkungen des eigenen Handelns** auf andere Personen hinterfragt wird. **HF: Hinterfragen** Erfüllt, wenn bei Nachfrage nach dem Verständnis von *Selbstreflexion* diese als das **Hinterfragen von Themen oder Dingen** verstanden wird, ohne den Prozess weiter auszuführen.	**N=1**

Verständnis Selbstreflexion (VS)	**VS-ET: Explizites Teilverständnis (ET)**	Der Begriff der *Selbstreflexion* liegt mit einer **informierten Definition** vor, welche **explizit einen Teilbereich der Grunddefinition** enthält.	IS-AL: "Selbstreflexion meint die introspektive Auseinandersetzung mit der eigenen Person zu einem Thema." (S3.F1-SS23-B-B3_5-XJ27BO, Pos. 2) IS-PK: "Das eigene Nachdenken über die eigene Persönlichkeit sollte dabei besonders im Vordergrund stehen." (S3.F2-SS23-B-B3_5-XF21FM, Pos. 2) IS-GE: "Die Fähigkeit meine eigenen Gedanken, Meinungen und Handlungen zu hinterfragen und einzuordnen." (S3.F2-SS23-B-B3_5-VM20MG, Pos. 2) IS-EM: "Die Reflexion der eigenen Gefühle [...]" (S3.F2-SS23-B-B3_5-BN10LB, Pos. 2) IS-WI: "Selbst zu reflektieren, was man weiß und was nicht." (S3.F1-SS23-B-B3_5-XJ34UB, Pos. 2) IS-AN: "Die Fähigkeit die eigenen Einstellungen und Meinungen zu reflektieren." (S3.F2-SS23-B-B3_5-GJ10JF, Pos. 2) IS-MW: "[...] in Beziehung zu setzen auf die Normen und Werte die man selbst pflegt oder die man von seiner Umgebung erfährt." (S3.F1-SS23-B-B3_5-CP17JN, Pos. 2) IS-FA: "Sich über seine eigenen Fähigkeiten klar zu werden (.)" (S3.F1-SS23-B-B3_5-IP17SE, Pos. 2) IS-EF: "Gedankliche Auseinandersetzungen mit eigenen Erfahrungen [...]" (S3.F1-SS23-B-B3_5-IB41VT, Pos. 2)	**Subkategorien:** **IS: Introspektion** **Innere Einsichtnahme, welche in folgende Kategorien weiter differenziert werden kann:** **IS-AL: Introspektion allgemein** Erfüllt, wenn der Begriff der *Selbstreflexion* als die **Fähigkeit zur Introspektion** benannt wird, ohne dies weiter auszuführen. **IS-PK: Persönlichkeit** Erfüllt, wenn der Begriff der *Selbstreflexion* als die Auseinandersetzung mit der eigenen **Persönlichkeit** benannt wird. **IS-GE: Gedanken** Erfüllt, wenn der Begriff der *Selbstreflexion* als die Auseinandersetzung mit den eigenen **Gedanken** benannt wird. **IS-EM: Emotionen** Erfüllt, wenn der Begriff der *Selbstreflexion* als die Auseinandersetzung mit den eigenen **Emotionen** benannt wird. **IS-WI: Wissen** Erfüllt, wenn der Begriff der *Selbstreflexion* als die Auseinandersetzung mit oder Bewertung des eigenen **Wissens** benannt wird. **IS-AN: Ansichten/Meinungen** Erfüllt, wenn der Begriff der *Selbstreflexion* als die Auseinandersetzung mit den eigenen **Ansichten oder der eigenen Meinung** benannt wird. **IS-MW: Moral/Werte** Erfüllt, wenn der Begriff der *Selbstreflexion* als die Auseinandersetzung mit den eigenen **Moral- oder Wertevorstellungen** benannt wird. **IS-FA: Fähigkeiten** Erfüllt, wenn der Begriff der *Selbstreflexion* als die Auseinandersetzung mit den eigenen Fähigkeiten oder die **Bewertung ebenjener** benannt wird. **IS-EF: Erfahrungen** Erfüllt, wenn der Begriff der *Selbstreflexion* als die Auseinandersetzung mit den eigenen **Erfahrungen** benannt wird.	**N=2**

Verständnis Selbstreflexion (VS)	**VS-ET: Explizites Teilverständnis (ET)**	Der Begriff der *Selbstreflexion* liegt mit einer **informierten Definition** vor, welche **explizit einen Teilbereich der Grunddefinition** enthält.	HA: "Das man sein eigenes Handeln und die Entscheidungen reflektiert." (S3.F2-SS23-B-B3_5-XJ34UB, Pos. 2) EX: "Das man bewusst wird, welche Einflüsse von außen auf einen wirken oder eingewirkt haben." (S3.F2-SS23-B-B3_5-XJ34UB, Pos. 2)	**HA: Überdenken von Handlungen** Erfüllt, wenn der Begriff der *Selbstreflexion* als die **ex post Bewertung eigenen Handelns** oder Verhaltens verstanden wird. **EX: Bewusstsein für externe Prägungen** Erfüllt, wenn der Begriff der *Selbstreflexion* als das **Hinterfragen von externen oder internen prägenden Faktoren** verstanden wird.	**N=2**

Tabelle A 12: Kodierleitfaden S3.F1&2 – Verständnis Kritisches Denken (VK)

S3.F1&2 - Verständnis Kritisch Denken (VK)					
VK: Welches Verständnis haben Studierende des Chemielehramtes von dem Begriff "Kritisches Denken"?					
Definition: Kritisches Denken (KD) bezeichnet eine sorgfältige und systematische sowie vor allem begründete Entscheidungsfindung bei der Auseinandersetzung der*s Denkenden mit einem Objekt oder Subjekt. Diese Auseinandersetzung beinhaltet die Evaluation und Anwendung des eigenen Wissens und der eigenen Fähigkeiten auf kognitiver Ebene sowie die Berücksichtigung der eigenen Emotionen, Werte, Haltung und Motivation auf affektiver Ebene. Es ermöglicht zudem eine konstante Selbstregulation und Anpassung, die eine Schleife zur*m Denkenden selbst zurückzieht und sie/ihn nicht nur in die Position der*s Bewertenden, sondern auch in die Position der*s zu Bewertenden versetzt (nach Rafolt, 2019; Paul & Elder, 2014, Facione, 1990).					
Kontexteinheit:	**Kodiereinheit:**	**Abstraktionsniveau:**	**Strukturdimension:**		**Auswertungseinheit:**
Die Antwort auf die zugehörige offene Fragestellung von VF2.	Klare, bedeutungstragende Elemente in der Antwort, wie Satzteile oder Wortzusammenhänge.	Generalisierbare Aussagen, welche das individuelle Verständnis der Begrifflichkeit des *Kritischen Denkens* deutlich machen. Mittleres Abstraktionsniveau.	Sowohl explizite Nennungen als auch implizite Hinweise, die das individuelle Verständnis der Begrifflichkeit des *Kritisches Denkens* verdeutlichen.		Alle Antworten auf die offene Fragestellung VF2 in den zugehörigen Fragebögen zu Studie 3.
Kategorie:		**Definition:**	**Ankerbeispiel:**	**Kodierregel:**	**Niveau:**
Verständnis von Kritischem Denken (VK)	**VK-NA: Naive Vorstellung von der Begriffichkeit des Kritischen Denkens (NA)**	*Kritisches Denken* kann **nicht definiert** werden oder es liegen diffuse Definitionsansätze vor.	ND: Kritisch über etwas nachdenken (S3.F2-SS23-B-B3_5-IB41VT, Pos. 3) HF: "Wenn man hinterfragen kann." (S3.F1-SS23-B-B3_5-TD35FO, Pos. 4) SK: "Nicht alles zu glauben, was man hört." (S3.F1-SS23-B-B3_5-XJ34UB, Pos. 3) SV: "Also ich übe sehr viel Selbstkritik so, das hättest du besser machen können oder das und das". (S2.I1-Wi22_23-B-B2_3-BN10LB, Pos. 60) TK: "Aber andere Menschen kritisiere ich nicht gerne oder mache es nicht häufig. Ich mache es nur, wenn ich es für angebracht oder sinnvoll erhalte. Also denke so, dass es Sinn bringt, die Person (unv.) zu kritisieren. Ich weiß nicht. Ich nehme meistens eher die Menschen so wie sie sind. Also das sind wirklich kritische Sachen, wie jemand übt Rassismus aus oder Ungleichheiten, dann sag ich da schon was." (S2.I1-Wi22_23-B-B2_3-BN10LB, Pos. 60) SD: "Selbstständig, ohne Leitung anderer, über ein Thema nachdenken [...]" (S3.F1-SS23-B-B3_5-EF28SH, Pos. 3)	**Subkategorien:** **ND: Nicht definierbar:** Erfüllt, wenn eine **diffuse oder keine Definition** von der Begrifflichkeit des *Kritischen Denkens* vorliegt. **HF: Hinterfragen:** Erfüllt, wenn unter der Begrifflichkeit des *Kritischen Denkens* das **Hinterfragen von Dingen und Themen** verstanden wird, wobei dies nicht weiter ausgeführt wird. **SK: Skepsis** Erfüllt, wenn unter der Begrifflichkeit des *Kritischen Denkens* verstanden wird, **nicht alles direkt zu glauben** oder **skeptisch zu sein**. **SV: Selbstverbesserung/Kritik:** Erfüllt, wenn unter der Begrifflichkeit des *Kritischen Denkens* ein **Prozess der Selbstverbesserung** verstanden wird, hervorgehend aus der **ex post Betrachtung** durchgeführter Handlungen. **TK: Tadeln bzw. Kritisieren:** Erfüllt, wenn unter der Begrifflichkeit des *Kritischen Denkens* das **Kritisieren** im Sinne von **Tadeln** anderer verstanden wird. **SD: Selbstständiges Denken** Erfüllt, wenn unter der Begrifflichkeit des *Kritischen Denkens* **eigenständiges Denken** verstanden wird.	**N=0**

Verständnis von Kritischem Denken (VK)	**VK-TG: Teilbereich der Grunddefinition (TG)**	Die Begrifflichkeit des *Kritischen Denkens* liegt mit einer informierten Definition vor, welche **explizit einen Teilbereich der Grunddefinition** enthält.	MP: Häufig wird versucht so viele Aspekte und Dimensionen wie möglich zu betrachten, um eine multiperspektive Sichtweise erhalten zu können. (S3.F2-SS23-B-B3_5-BN10LB, Pos. 3) SR: Eigene Meinung und Werte in Frage stellen, wenn nötig. (S3.F1-SS23-B-B3_5-CP17JN), Pos. 3) WI: "Darunter verstehe ich das kritische Nachfragen von Aussagen. Inwiefern können diese verifiziert oder falsifiziert werden." (S3.F1-SS23-B-B3_5-BN27DI, Pos. 3) EP: "Fähigkeit, Informationsquellen auf Glaubwürdigkeit zu prüfen und zu beurteilen [...]" (S3.F2-SS23-B-B3_5-SJ16DI, Pos. 3)	**<u>Subkategorien:</u>** **MP: Multiperspektivität** Erfüllt, wenn unter der Begrifflichkeit des *Kritischen Denkens* eine **multiperspektivische Betrachtung** verstanden wird. **SR: Selbstreflexion** Erfüllt, wenn unter der Begrifflichkeit des *Kritischen Denkens* das **Hinterfragen von sich selbst oder den eigenen Prägungen** im Sinne einer **Selbstreflexion** verstanden wird. **WI: Wissen** Erfüllt, wenn unter der Begrifflichkeit des *Kritischen Denkens* das **Hinterfragen von externen Informationen oder eigenem Wissen** verstanden wird. Ebenfalls erfüllt, wenn deutlich wird, dass **faktenorientiertes Wissen** zum Kritischen Denken gehört. **EP: Epistemologisches Wissen** Erfüllt, unter der Begrifflichkeit des *Kritischen Denkens* das **Hinterfragen und Bewerten von Quellen** verstanden wird.	**N=1**
	VK-MT: Zusammenspiel mehrerer Teilbereiche der Grunddefinition (MT)	Die Begrifflichkeit des *Kritischen Denkens* wird als **Zusammenspiel** von Selbstreflexion, epistemologischem Verständnis und Fachwissen verstanden.		Erfüllt, wenn bei der Definition der Begrifflichkeit des *Kritischen Denkens* dieses als **Zusammenspiel von mindestens zwei der Teilbereiche der Grunddefinition** verstanden wird: Fachwissen, epistemologisches Verständnis (Quellenbewertung) und Selbstreflexion. Immer dann kodiert, wenn mindestens zwei der Subkategorien von TG kodiert wurden.	**N=2**

Tabelle A 13: Kodierleitfaden S3.F1&2 – Beeinflussende Faktoren bei der Bewertung (BF)

S3.F1&2 - Beeinflussende Faktoren bei der Bewertung (BF)						
BF: Welche Faktoren werden von Chemielehramtsstudierenden als beeinflussend bei der Bewertung von Themen mit naturwissenschaftlichem Bezug wahrgenommen?						
Kontexteinheit:	**Kodiereinheit:**		**Abstraktionsniveau:**	**Strukturdimension:**		**Auswertungseinheit:**
Die Antwort auf die zugehörige offene Fragestellung von EF1 und VF3	Klare, bedeutungstragende Elemente in der Antwort, wie Satzteile oder Wortzusammenhänge.		Generalisierbare Aussagen zu Aspekten oder Faktoren, die als beeinflussend im Bewertungskontext wahrgenommen werden. Hohes Abstraktionsniveau.	Sowohl explizite Nennungen als auch implizite Hinweise, die deutlich machen, welche Aspekte oder Faktoren im Bewertungskontext als beeinflussend wahrgenommen werden.		Alle Antworten auf die offene Fragestellung EF1 und VF3 in den zugehörigen Fragebögen zu Studie 3.
Kategorie:			**Definition:**	**Ankerbeispiel:**	**Kodierregel:**	**Kodierung:**
SR: Selbstreflexive Aspekte	**SR-IF: Interne Faktoren**	**BF-SR-IF-FW: Fachwissen**	Das **Vorwissen oder die fachliche Grundlage** werden als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"[…]wie viel und was genau, ich über das Thema schon weiß (.)" (S3.F2-SS23-B-B3_5-XF21FM, Pos. 2)	Erfüllt, wenn deutlich wird, dass das **Vorwissen** als beeinflussender Aspekt im Bewertungskontext wahrgenommen wird..	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-IF-PR: Persönliche Relevanz/ Persönlicher Bezug	Die **persönliche Relevanz** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen	"Relevanz, Wichtigkeit, Präsenz in meinem Leben […]" (S3.F1-SS23-B-B3_5-XF21FM, Pos. 2)	Erfüllt, wenn deutlich wird, dass die **persönliche Relevanz** als beeinflussenden Aspekt im Bewertungskontext wahrgenommen wird. Dies schließt das **Abwägen** von **Vor- und Nachteilen** für die eigene Person ein, ebenso wie die **Bewertung des eigenen Möglichkeitsrahmens** oder der **wahrgenommenen Notwendigkeit**.	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-IF-MW: Moral, Werte und Normen	**Moraltvorstellungen und Werte** werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen	"Meine Normen, Werte […]" (S3.F2-SS23-B-B3_5-SV32BT, Pos. 4)	Erfüllt, wenn deutlich wird, dass eigene **Moraltvorstellungen und Werte** als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden. Auch erfüllt, wenn die **persönliche Identifikation** angeführt wird.	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-IF-HA: Handlungen	**Vorangegangene Handlungen** werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen	"Meine Normen, Werte […]" (S3.F2-SS23-B-B3_5-SV32BT, Pos. 4)	Erfüllt, wenn deutlich wird, dass eigene vorangegangene Handlungen als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden.	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-IF-EM: Emotionen	**Emotionen** werden als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"[…] meine Gefühlswelt zum Thema (.)" (S3.F1-SS23-B-B3_5-GJ10JF, Pos. 2)	Erfüllt, wenn deutlich wird, dass eigene **Emotionen** als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden.	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-IF-EF: Erfahrungen	**Gemachte Erfahrungen** werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen.	"[…] Persönliche Erfahrungen (.)" (S3.F2-SS23-B-B3_5-VM20MG, Pos. 2)	Erfüllt, wenn deutlich wird, dass **gemachten Erfahrungen** als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-IF-IN: Interesse	**Das persönliche Interesse** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	" Wenn das Thema mich interessiert […]" (S3.F2-SS23-B-B3_5-HM28EP, Pos. 2)	Erfüllt, wenn deutlich wird, dass das **persönliche (Des-)Interesse** als beeinflussender Aspekt im Bewertungskontext wahrgenommen wird..	Nicht erfüllt: 0 Erfüllt: 1

SR: Selbstreflexive Aspekte	**SR-IF: Interne Faktoren**	**BF-SR-IF-BV: Befangenheit, Vorurteile**	**Die Befangenheit oder die Meinung** werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen.	"[…] meine subjektive Meinung zu dem Sachverhalt." (S3.F1-SS23-B-B3_5-WJ32FE, Pos. 4)	Erfüllt, wenn deutlich wird, dass die eigene **Befangenheit oder Meinung** als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden. Ebenfalls erfüllt, wenn **Antipathie oder Sympathie** angeführt werden.	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-IF-PK: Persönlichkeit	Die **Persönlichkeit** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"[…] mein Idealismus […] (S3.F2-SS23-B-B3_5-SJ16DI, Pos. 4)	Erfüllt, wenn deutlich wird, dass die eigene **Persönlichkeit** als beeinflussender Aspekt im Bewertungskontext wahrgenommen wird.	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-EF-KV: Kognitive Verzerrungen	Eine **verzerrte Wahrnehmung** durch kognitive Verzerrungen wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"[…] kognitive Verzerrungen (.)" (S3.F2-SS23-B-B3_5-EF28SH, Pos. 4)	Erfüllt, wenn deutlich wird, dass die **verzerrte Wahrnehmung** durch kognitive Verzerrungen als beeinflussender Aspekt im Bewertungskontext wahrgenommen wird.	Nicht erfüllt: 0 Erfüllt: 1
SR: Selbstreflexive Aspekte	**SR-EF: Externe Faktoren**	**BF-SR-EF-SN: Soziale Norm/Umfeld**	Das **persönliche Umfeld** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Einerseits hat mein Umfeld ein enormen Einfluss auf mein Verhalten." S3.F2-SS23-B-B3_5-BN10LB, Pos. 4)	Erfüllt, wenn deutlich wird, dass das **persönliches Umfeld** als beeinflussender Aspekt im Bewertungskontext wahrgenommen wird. Hierunter fallen auch Nennungen der **sozialen Erwünschtheit**, **Prägungen durch die Familie und Freude** oder die **Erziehung**.	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-EF-ME: Medien	**Medien** werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen.	"Auf der einen Seite würde ich sagen, dass mich Social Media beeinflussen würde, weil wir heute in einer so digital präsenten Welt leben, in der man z.B. auf Informationen auch erst Recht aus dem Internet stößt." (S3.F1-SS23-B-B3_5-XF21FM, Pos. 4)	Erfüllt, wenn deutlich wird, dass **Medien** als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden. Hierzu zählen Einflüsse aus **konservativen Medien** als auch **soziale Medien** und das **Internet**.	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-EF-BI: Studium/Bildung	Die eigene **Vorbildung** oder das **Studium** werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen.	"Mein Lehramtsstudium, besonders die Fächer Geographie und Chemie." (S3.F2-SS23-B-B3_5-GJ10JF, Pos. 4)	Erfüllt, wenn deutlich wird, dass die eigene **Vorbildung** oder das **Studium** als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden.	Nicht erfüllt: 0 Erfüllt: 1
		BF-SR-EF-PO: Politische Position	Die **politische Position** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Meine politische Positionierung […]" (S3.F2-SS23-B-B3_5-SJ16DI, Pos. 4)	Erfüllt, wenn deutlich wird, dass die eigene **politische Position** als beeinflussender Aspekt im Bewertungskontext wahrgenommen wird.	Nicht erfüllt: 0 Erfüllt: 11
		BF-SR-EF-RL: Religion	**Religiöse Anschauungen** werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen.	"Der Glaube. Kann das übereinstimmen mit dem was ich aus der Bibel kenne." (S3.F1-SS23-B-B3_5-XJ34UB, Pos. 4)	Erfüllt, wenn deutlich wird, dass **religiöse Anschauungen** als beeinflussenden Aspekte im Bewertungskontext wahrgenommen werden.	Nicht erfüllt: 0 Erfüllt: 1

Kategorie:		Definition:	Ankerbeispiel:	Kodierregel:	Kodierung:
EP: Epistemologische Aspekte	**BF-EP-EX: Autoritäten/ Expertise**	Die **Expertise von Personen** oder Institution sowie Autoritäten wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Autoritäten, ich lasse mich besonders dazu verleiten, wenn ein Wissenschaftler/eine Wissenschaftlerin, der/die seriös wirkt etwas sagt, die Sache nicht zu hinterfragen, obwohl ich es eigentlich trotzdem sollte." (S3.F1-SS23-B-B3_5-EF28SH, Pos. 4)	Erfüllt, wenn deutlich wird, dass die **Autorität und Expertise** der*s Autor*in als relevante Aspekte bei der Einschätzung der Vertrauenswürdigkeit einer Quelle wahrgenommen werden.	Nicht erfüllt: 0 Erfüllt: 1
	BF-EP-AK: Aktualität	Die **Aktualität** einer Information wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Zum einen die Aktualität […]" (S3.F2-SS23-B-B3_5-XF21FM, Pos. 2)	Erfüllt, wenn deutlich wird, dass die **Aktualität** der Veröffentlichung als relevanter Aspekt bei der Einschätzung der Vertrauenswürdigkeit einer Quelle wahrgenommen wird.	Nicht erfüllt: 0 Erfüllt: 1
	BF-EP-IK: Interessenskonflikte	**Interessenskonflikte** werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen.	"Es gibt ja liberale, konservative Zeitungen, die dann dementsprechend auch Meinungen so oder so vertreten, oder auch bei gewissen Autoren ist es bekannt, dass die absolut diese Stellung immer einnehmen [...]" (S3.I2-SS23-A-B3_5_WJ32FE, Pos. 25)	Erfüllt, wenn deutlich wird, dass **etwaige Interessenskonflikte** oder Subjektivität (Meinungen) als relevante Aspekte bei der Einschätzung der Vertrauenswürdigkeit einer Quelle wahrgenommen werden. Auch die umgekehrte Formulierung der gewünschten Unabhängigkeit von Quellen oder Berichterstattern werden hier kodiert.	Nicht erfüllt: 0 Erfüllt: 1
	BF-EP-PR: Präsentation	Der **Stil** oder die **Präsentation** von Informationen werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen.	"Wird es mit der nötigen Fachsprache auch vermittelt?" (BN27DI-Pre, Pos. 2) "[…] Art und Weise wie es vermittelt wird (.)" (S3.F2-SS23-B-B3_5-XF36JE, Pos. 2)	Erfüllt, wenn deutlich wird, dass die **Präsentation oder Sprache und Stil** der Quelle als relevante Aspekte bei der Einschätzung der Vertrauenswürdigkeit einer Quelle wahrgenommen werden. Auch erfüllt, wenn Schlüssigkeit und **Logik der Argumente** angeführt werden.	Nicht erfüllt: 0 Erfüllt: 1
	BF-EP-HK: Herkunft der Information	Die **Herkunft der Information** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"[…]die Art der Quelle der Informationen (wie seriös die Quelle beurteilt wird (.)" (S3.F2-SS23-B-B3_5-WJ32FE, Pos. 4)	Erfüllt, wenn deutlich wird, dass die **Herkunft oder Veröffentlichungsplattform** sowie deren Bekanntheit als relevante Aspekte bei der Einschätzung der Vertrauenswürdigkeit einer Quelle wahrgenommen werden.	Nicht erfüllt: 0 Erfüllt: 1
	BF-EP-ME: Methodik	Die **Methodik, mit der Informationen gewonnen werden,** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Methodik (.)" (S3.F1-SS23-B-B3_5-LB32QI, Pos. 4)	Erfüllt, wenn deutlich wird, dass die **Methodik** und das Forschungsdesign als relevante Aspekte bei der Einschätzung der Vertrauenswürdigkeit einer Quelle wahrgenommen werden.	Nicht erfüllt: 0 Erfüllt: 1
	BF-EP-KO: Konsistenzprüfung	Die **Übereinstimmung mehrerer Quellen** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Aber ich glaube, ich würde schon darauf achten, verschiedene Quellen hin, also bei der Recherche hinzuzuziehen. Und nicht nur die Windkraftexperten oder nicht nur die Solarenergieexperten, oder Befürworter, also die Quellen zu nutzen, sondern ich glaube, ich würde schon darauf achten, dass es so breit gefächert ist." (S3.I1-SS23-A-B3_5-SJ16DI, Pos. 44)	Erfüllt, wenn deutlich wird, dass die **Übereinstimmung mit anderen Quellen** als relevanter Aspekt bei der Einschätzung der Vertrauenswürdigkeit einer Quelle wahrgenommen wird. Ebenfalls erfüllt, wenn **andere Standpunkte bei der Bewertung hinzugezogen** werden, um eine bessere Aussagekraft zu erreichen.	Nicht erfüllt: 0 Erfüllt: 1

EP: Epistemologische Aspekte	**BF-EP- TR: Transparenz**	Die **transparente Angabe von Quellen** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Also genau, ich würde natürlich auch gucken, wie die Informationen aufbereitet sind, also wie, ob Studien mit Namen angegeben sind oder vielleicht verlinkt sind oder wie genau die Quellen angegeben sind und welche Quellen angegeben sind, ob das jetzt wissenschaftliche Publikationen sind oder andere Websites oder so was." (S3.I2-SS23-B-B3_5_SJ16DI, Pos. 17)	Erfüllt, wenn deutlich wird, dass eine offene und **transparente Angabe von (wissenschaftlichen) (Erst-)Quellen** oder Herkunftsinformationen als relevanter Aspekt bei der Einschätzung der Vertrauenswürdigkeit einer Quelle wahrgenommen wird.	Nicht erfüllt: 0 Erfüllt: 1
	BF-EP-SE: Soziokulturelle Eingebundenheit	Der **kulturellen Einfluss** im Sinne einer soziokulturellen Eingebundenheit wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Soziokulturelle Umstände […]" (S3.F1-SS23-B-B3_5-VM20MG, Pos. 2)	Erfüllt, wenn deutlich wird, dass die **soziokulturelle Eingebundenheit** der Erkenntnisse, der*s Autorin*s oder der Plattform als relevanter Aspekt bei der Einschätzung der Vertrauenswürdigkeit einer Quelle wahrgenommen wird.	Nicht erfüllt: 0 Erfüllt: 1

Kategorie:			**Definition:**	**Ankerbeispiel:**	**Kodierregel:**	**Kodierung:**
WI: Wissensbezogene Aspekte	**Allgemeine Aspekte**	**BF-WI-WS: Wissenschaftliche Standards**	Die Einhaltung **wissenschaftlicher Standards** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Ob diese empirisch belegt sind (.)" (S3.F2-SS23-B-B3_5-HM28EP, Pos. 4)	Erfüllt, wenn deutlich wird, dass **wissenschaftliche Standards**, wie eine faktenbasierte Betrachtung, empirische Grundlage, Objektivität oder Vergleichbares, als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden.	Nicht erfüllt: 0 Erfüllt: 1
	Spezifische Aspekte	**BF-WI-EF: Effizienz**	Die **Effizienz** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Was es einfach für Zahlen gibt und für Werte, also wie effizient diese Energiegewinnung ist." (S3.I1-SS23-A-B3_5-SJ16DI, Pos. 42)	Erfüllt, wenn deutlich wird, dass die **Effizienz** als beeinflussender Aspekt im Bewertungskontext wahrgenommen wird.	Nicht erfüllt: 0 Erfüllt: 1
		BF-WI-KA: Kostenabwägung	Die **Kosten** oder das Kosten-Nutzen-Verhältnis werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen.	"Die Zusammensetzung der Lithiumakkus und auch wie, ich sage mal, wie, wie preisgünstig die auch produziert werden können und was E-Fuels und Wasserstoff angeht, […]" (S3.I1-SS23-B-B3_5-WJ32FE, Pos. 42)	Erfüllt, wenn deutlich wird, dass die **Kosten** oder das Kosten-Nutzen-Verhältnis als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden.	Nicht erfüllt: 0 Erfüllt: 1
		BF-WI-NA: Nachhaltigkeit und Auswirkungen	Die **Auswirkungen** und/oder **Nachhaltigkeit** werden als beeinflussende Aspekte im Bewertungskontext wahrgenommen.	"Aber so den Punkt der Nachhaltigkeit, würde ich auf jeden Fall sagen, wäre bei mir so an oberster Stelle und dann würde ich halt schauen okay, was ist nachhaltiger, was ist dafür sinnvoller? Und was sind jetzt die Negativaspekte, auch wenn das Ding jetzt nachhaltiger ist und man die gegeneinander abwägt?" (S3.I2-SS23-B-B3_5_XJ36JE, Pos. 21)	Erfüllt, wenn deutlich wird, dass die **Auswirkungen** auf gesundheitlicher sozialer oder ökologischer Nachhaltigkeitebene als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden.	Nicht erfüllt: 0 Erfüllt: 1

WI: Wissensbezogene Aspekte	**Spezifische Aspekte**	**BF-WI-FT: Funktionsweise**	Die **Funktionsweise** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"Dann noch von irgendwie den E-Autoherstellern wie das Ganze, wie der Herstellungsprozess da vonstattengeht, was da alles berücksichtigt werden muss, wie das Ganze funktioniert. Sich einfach mit der Theorie befassen, die hinter so einem E-Auto steckt oder hinter so ein Brennstoffzellenauto steckt, damit man einfach mal weiß, nicht nur, dass ich das Auto fahren kann, sondern dass ich auch weiß, wie es funktioniert." (S3.I2-SS23-B-B3_5-XJ27BO, Pos. 23)	Erfüllt, wenn deutlich wird, dass das Verstehen der **Funktionsweise** oder der zugrundelegenden **Technik** als als beeinflussender Aspekt im Bewertungskontext wahrgenommen wird. Auch erfüllt, wenn Definitionen für ein Grundverständnis erfragt werden.	Nicht erfüllt: 0 Erfüllt: 1
		BF-WI-UM: Umsetzbarkeit	Die **Umsetzbarkeit** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"[…] vielleicht ist jetzt das eine super umweltfreundlich, aber ist das überhaupt realistisch? Oder ist das vielleicht einfach nur so ein Wunschdenken? Und kann das überhaupt in der Realität funktionieren?" (S3.I2-SS23-A-B3_5_EF28SH, Pos. 23)	Erfüllt, wenn deutlich wird, dass die **Umsetzbarkeit** als beeinflussender Aspekt im Bewertungskontext wahrgenommen wird. Auch erfüllt, wenn Herausforderungen oder Barrieren für die Umsetzung diskutiert werden.	Nicht erfüllt: 0 Erfüllt: 1
		BF-WI-RH: Rohstoffherkunft	Die **Herkunft von Rohstoffen** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"[...] wo kommen die Rohstoffe her, die da eingebaut werden? Wenn man jetzt irgendwie alternative Akkus hätte, die man irgendwie in Deutschland produzieren könnte, mit Rohstoffen, Ressourcen, die durchaus mehr zur Verfügung stehen würden als Lithium, dann hätte ich gesagt auf jeden Fall E-Auto." (S3.I2-SS23-B-B3_5-XJ27BO, Pos. 13)	Erfüllt, wenn deutlich wird, dass die **Herkunft von Rohstoffen oder Bauteilen** als beeinflussender Aspekt im Bewertungskontext wahrgenommen wird.	Nicht erfüllt: 0 Erfüllt: 1
		BF-WI-NO: Notwendigkeit	Die **Notwendigkeit** wird als beeinflussender Aspekt im Bewertungskontext wahrgenommen.	"[…] inwieweit welche Energien am intensivsten gebraucht werden, bezogen auf Energiepolitik und inwieweit erneuerbare Energien ausgebaut wurden und noch werden." (S2.I1-Wi22_23-A2-B2_3-XJ27BO, Pos. 30)	Erfüllt, wenn deutlich wird, dass der **Bedarf oder die Notwendigkeit** als beeinflussende Aspekte im Bewertungskontext wahrgenommen werden. Auch erfüllt, wenn Alternativen erfragt werden, da dann eine implizite Notwendigkeitsabfrage vorliegt.	Nicht erfüllt: 0 Erfüllt: 1

Danksagung

An dieser Stelle möchte ich mich bei all den Menschen bedanken, die zum Gelingen dieser Arbeit beigetragen und mich auf meinem Weg unterstützt haben.

Allen voran danke ich meiner Doktormutter, Frau Prof. Dr. Christiane Reiners, die es mir ermöglicht hat, das Thema des Kritischen Denkens weiterzuverfolgen, und einen Möglichkeitsraum geschaffen hat, in dem ich meine Perspektiven erweitern konnte. Für ihr unermüdliches Engagement, die vielen tiefgreifenden und konstruktiven Gespräche sowie die Ermunterungen im Prozess dieser Arbeit möchte ich mich ganz herzlich bei ihr bedanken. Besonders bedanken möchte ich mich zudem bei Frau Prof. Dr. Katharina Groß für die Übernahme des Zweitgutachtens, die Unterstützung im Institut sowie für die Chance, wertvolle internationale Erfahrungen im Forschungsprozess sammeln zu können. Darüber hinaus danke ich Herrn Prof. Dr. André Bresges ganz herzlich für die Übernahme des Vorsitzes der Prüfungskommission sowie Herrn. Dr. Markus Krüger für den Prüfungsbeisitz.

Des Weiteren möchte ich mich bei allen aktuellen und ehemaligen Mitarbeiter*innen des Instituts für Chemiedidaktik bedanken. Ich danke Frau Blockhaus für ihre immer freundliche Art im Sekretariat sowie Frau Heinen-Ebert und Frau Gretz für die Hilfsbereitschaft in allen Zwischenfragen. Ferner danke ich meinen Mitdoktorand*innen für die gemeinsame Zeit und schöne Zusammenarbeit. Ausdrücklich möchte ich mich an dieser Stelle bei Herrn Tom Konrad Anton für den fachlichen Austausch und die menschlichen Gespräche bedanken. An dieser Stelle möchte ich zudem den studentischen Hilfskräften für die gemeinsamen Kaffeepausen und die Unterstützung bei der Literaturrecherche meinen Dank aussprechen, insbesondere Herrn Tim Emmerich für seine wertvolle Hilfe. Ein weiterer Dank gebührt den Studierenden, die durch ihre Mitwirkung an den empirischen Studien diese Arbeit erst möglich gemacht haben.

Neben den fachlichen Wegbegleiter*innen möchte ich mich außerdem von ganzem Herzen bei meiner Mutter Renate Albrecht bedanken, die mich auf meinem Weg immer unterstützt hat, und bei meinem Vater Ben Albrecht, der mir die Ressourcen mitgab, aus denen ich heute schöpfe. Ebenfalls danke ich meiner (Schwieger-)Familie und meinen Freunden für die emotionale Unterstützung, ihr Interesse und die gelegentlich nötige Zerstreuung. Abschließend danke ich meiner zukünftigen Ehefrau Luise Maul für ihre bedingungslose Unterstützung, für ihren Rat und ihren Zuspruch

in allen Phasen dieser Arbeit. Ohne ihre Geduld, ihre Ermutigungen und nicht zuletzt ihre wertvolle Lektoratsarbeit wäre diese Dissertation in dieser Form nicht möglich gewesen.

Teilpublikationen

Albrecht, L. B., & Reiners, C. S. (2025). Kritisches Denken in den Naturwissenschaften: Analyse und Konsequenzen. In H. van Vorst (Hrsg.), *Entdecken, lehren und forschen im Schülerlabor, GDCP-Tagungsband* (Bd. 45, S. 441–444). GDCP.

Albrecht, L. B. (2024). Fostering Critical Thinking in Science Education: Definitions, Assessments, and Implications for Sustainable Decision-Making. *Book of Abstracts of ECRICE 2024, the 16 European Conference on Research in Chemical Education*, *16*, 88–89.

Albrecht, L. B., & Reiners, C. S. (2024). Kritisches Denken als Schlüssel zur naturwissenschaftlichen Bildung. In H. van Vorst (Hrsg.), *Frühe naturwissenschaftliche Bildung, GDCP-Tagungsband* (Bd. 44, S. 334–337). GDCP.

Bisher erschienene Bände der Reihe
Studien zum Physik- und Chemielernen

ISSN 1614-8967

Vollständige Übersicht auf unserer Website

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

Aktuelle Bände

310 Wolfgang Becker (2021): Auswirkungen unterschiedlicher experimenteller Repräsentationen auf den Kenntnisstand bei Grundschulkindern
ISBN 978-3-8325-5255-8 50.00 EUR

311 Marvin Rost (2021): Modelle als Mittel der Erkenntnisgewinnung im Chemieunterricht der Sekundarstufe I. Entwicklung und quantitative Dimensionalitätsanalyse eines Testinstruments aus epistemologischer Perspektive
ISBN 978-3-8325-5256-5 44.00 EUR (open access)

312 Christina Kobl (2021): Förderung und Erfassung der Reflexionskompetenz im Fach Chemie
ISBN 978-3-8325-5259-6 41.00 EUR (open access)

313 Ann-Kathrin Beretz (2021): Diagnostische Prozesse von Studierenden des Lehramts. eine Videostudie in den Fächern Physik und Mathematik
ISBN 978-3-8325-5288-6 45.00 EUR (open access)

314 Judith Breuer (2021): Implementierung fachdidaktischer Innovationen durch das Angebot materialgestützter Unterrichtskonzeptionen. Fallanalysen zum Nutzungsverhalten von Lehrkräften am Beispiel des Münchener Lehrgangs zur Quantenmechanik
ISBN 978-3-8325-5293-0 50.50 EUR (open access)

315 Michaela Oettle (2021): Modellierung des Fachwissens von Lehrkräften in der Teilchenphysik. Eine Delphi-Studie
ISBN 978-3-8325-5305-0 57.50 EUR (open access)

316 Volker Brüggemann (2021): Entwicklung und Pilotierung eines adaptiven Multistage-Tests zur Kompetenzerfassung im Bereich naturwissenschaftlichen Denkens
ISBN 978-3-8325-5331-9 40.00 EUR (open access)

317 Stefan Müller (2021): Die Vorläufigkeit und soziokulturelle Eingebundenheit naturwissenschaftlicher Erkenntnisse. Kritische Reflexion, empirische Befunde und fachdidaktische Konsequenzen für die Chemielehrer*innenbildung
ISBN 978-3-8325-5343-2 63.00 EUR

318 Laurence Müller (2021): Alltagsentscheidungen für den Chemieunterricht erkennen und Entscheidungsprozesse explorativ begleiten
ISBN 978-3-8325-5379-1 59.00 EUR

319 Lars Ehlert (2021): Entwicklung und Evaluation einer Lehrkräftefortbildung zur Planung von selbstgesteuerten Experimenten
ISBN 978-3-8325-5393-7 41.50 EUR (open access)

320 Florian Seiler (2021): Entwicklung und Evaluation eines Seminarkonzepts zur Förderung der experimentellen Planungskompetenz von Lehramtsstudierenden im Fach Chemie
ISBN 978-3-8325-5397-5 47.50 EUR (open access)

321 Nadine Boele (2021): Entwicklung eines Messinstruments zur Erfassung der professionellen Unterrichtswahrnehmung von (angehenden) Chemielehrkräften hinsichtlich der Lernunterstützung
ISBN 978-3-8325-5402-6 46.50 EUR

322 Franziska Zimmermann (2022): Entwicklung und Evaluation digitalisierungsbezogener Kompetenzen von angehenden Chemielehrkräften
ISBN 978-3-8325-5410-1 49.50 EUR

323 Lars-Frederik Weiß (2021): Der Flipped Classroom in der Physik-Lehre. Empirische Untersuchungen in Schule und Hochschule
ISBN 978-3-8325-5418-7 51.00 EUR

324 Tilmann Steinmetz (2021): Kumulatives Lehren und Lernen im Lehramtsstudium Physik. Theorie und Evaluation eines Lehrkonzepts
ISBN 978-3-8325-5421-7 51.50 EUR

325 Kübra Nur Celik (2022): Entwicklung von chemischem Fachwissen in der Sekundarstufe I. Validierung einer Learning Progression für die Basiskonzepte „Struktur der Materie“, „Chemische Reaktion“ und „Energie“ im Kompetenzbereich „Umgang mit Fachwissen“
ISBN 978-3-8325-5431-6 55.00 EUR

326 Matthias Ungermann (2022): Förderung des Verständnisses von Nature of Science und der experimentellen Kompetenz im Schüler*innen-Labor Physik in Abgrenzung zum Regelunterricht
ISBN 978-3-8325-5442-2 55.50 EUR

327 Christoph Hoyer (2022): Multimedial unterstütztes Experimentieren im webbasierten Labor zur Messung, Visualisierung und Analyse des Feldes eines Permanentmagneten
ISBN 978-3-8325-5453-8 45.00 EUR

328 Tobias Schüttler (2022): Schülerlabore als interessefördernde authentische Lernorte für den naturwissenschaftlichen Unterricht nutzen
ISBN 978-3-8325-5454-5 50.50 EUR

329 Christopher Kurth (2022): Die Kompetenz von Studierenden, Schülerschwierigkeiten beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-5457-6 58.50 EUR

330 Dagmar Michna (2022): Inklusiver Anfangsunterricht Chemie. Entwicklung und Evaluation einer Unterrichtseinheit zur Einführung der chemischen Reaktion
ISBN 978-3-8325-5463-7 49.50 EUR

331 Marco Seiter (2022): Die Bedeutung der Elementarisierung für den Erfolg von Mechanikunterricht in der Sekundarstufe I
ISBN 978-3-8325-5471-2 66.00 EUR

332 Jörn Hägele (2022): Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten. Videobasierte Analysen zu Aktivitäten und Vorstellungen von Schülerinnen und Schülern der gymnasialen Oberstufe bei der Bearbeitung von fachmethodischer Instruktion
ISBN 978-3-8325-5476-7 56.50 EUR (open access)

333 Erik Heine (2022): Wissenschaftliche Kontroversen im Physikunterricht. Explorationsstudie zum Umgang von Physiklehrkräften und Physiklehramtsstudierenden mit einer wissenschaftlichen Kontroverse am Beispiel der Masse in der Speziellen Relativitätstheorie
ISBN 978-3-8325-5478-1 48.50 EUR (open access)

334 Simon Goertz (2022): Module und Lernzirkel der Plattform FLexKom zur Förderung experimenteller Kompetenzen in der Schulpraxis. Verlauf und Ergebnisse einer Design-Based Research Studie
ISBN 978-3-8325-5494-1 66.50 EUR

335 Christina Toschka (2022): Lernen mit Modellexperimenten. Empirische Untersuchung der Wahrnehmung und des Denkens in Analogien beim Umgang mit Modellexperimenten
ISBN 978-3-8325-5495-8 50.00 EUR (open access)

336 Alina Behrendt (2022): Chemiebezogene Kompetenzen in der Übergangsphase zwischen dem Sachunterricht der Primarstufe und dem Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-5498-9 40.50 EUR (open access)

337 Manuel Daiber (2022): Entwicklung eines Lehrkonzepts für eine elementare Quantenmechanik. Formuliert mit In-Out Symbolen
ISBN 978-3-8325-5507-8 48.50 EUR

338 Felix Pawlak (2022): Das Gemeinsame Experimentieren (an-)leiten. Eine qualitative Studie zum chemiespezifischen Classroom-Management
ISBN 978-3-8325-5508-5 46.50 EUR

339 Liza Dopatka (2022): Konzeption und Evaluation eines kontextstrukturierten Unterrichtskonzeptes für den Anfangs-Elektrizitätslehreunterricht
ISBN 978-3-8325-5514-6 69.50 EUR

340 Arne Bewersdorff (2022): Untersuchung der Effektivität zweier Fortbildungsformate zum Experimentieren mit dem Fokus auf das Unterrichtshandeln
ISBN 978-3-8325-5522-1 39.00 EUR (open access)

341 Thomas Christoph Münster (2022): Wie diagnostizieren Studierende des Lehramtes physikbezogene Lernprozesse von Schüler*innen?. Eine Videostudie zur Mechanik
ISBN 978-3-8325-5534-4 44.50 EUR (open access)

342 Ines Komor (2022): Förderung des symbolisch-mathematischen Modellverständnisses in der Physikalischen Chemie
ISBN 978-3-8325-5546-7 46.50 EUR

343 Verena Petermann (2022): Überzeugungen von Lehrkräften zum Lehren und Lernen von Fachinhalten und Fachmethoden und deren Beziehung zu unterrichtsnahem Handeln
ISBN 978-3-8325-5545-0 47.00 EUR (open access)

344 Jana Heinze (2022): Einfluss der sprachlichen Konzeption auf die Einschätzung der Qualität instruktionaler Unterrichtserklärungen im Fach Physik
ISBN 978-3-8325-5553-5 42.00 EUR (open access)

345 Jannis Weber (2022): Mathematische Modellbildung und Videoanalyse zum Lernen der Newtonschen Dynamik im Vergleich
ISBN 978-3-8325-5566-5 68.00 EUR (open access)

346 Fabian Sterzing (2022): Zur Lernwirksamkeit von Erklärvideos in der Physik. Eine Untersuchung in Abhängigkeit von ihrer fachdidaktischen Qualität und ihrem Einbettungsformat
ISBN 978-3-8325-5576-4 52.00 EUR (open access)

347 Lars Greitemann (2022): Wirkung des Tablet-Einsatzes im Chemieunterricht der Sekundarstufe I unter besonderer Berücksichtigung von Wissensvermittlung und Wissenssicherung
ISBN 978-3-8325-5580-1 50.00 EUR

348 Fabian Poensgen (2022): Diagnose experimenteller Kompetenzen in der laborpraktischen Chemielehrer*innenbildung
ISBN 978-3-8325-5587-0 48.00 EUR

349 William Lindlahr (2023): Virtual-Reality-Experimente. Entwicklung und Evaluation eines Konzepts für den forschend-entwickelnden Physikunterricht mit digitalen Medien
ISBN 978-3-8325-5595-5 49.00 EUR

350 Bert Schlüter (2023): Teilnahmemotivation und situationales Interesse von Kindern und Eltern im experimentellen Lernsetting KEMIE
ISBN 978-3-8325-5598-6 43.00 EUR

351 Katharina Nave (2023): Charakterisierung situativer mentaler Modellkomponenten in der Chemie und die Bildung von Hypothesen. Eine qualitative Studie zur Operationalisierung mentaler Modell-komponenten für den Fachbereich Chemie
ISBN 978-3-8325-5599-3 43.00 EUR

352 Anna B. Bauer (2023): Experimentelle Kompetenz Physikstudierender. Entwicklung und erste Erprobung eines performanzorientierten Kompetenzstrukturmodells unter Nutzung qualitativer Methoden
ISBN 978-3-8325-5625-9 47.00 EUR (open access)

353 Jan Schröder (2023): Entwicklung eines Performanztests zur Messung der Fähigkeit zur Unterrichtsplanung bei Lehramtsstudierenden im Fach Physik
ISBN 978-3-8325-5655-6 46.50 EUR

354 Susanne Gerlach (2023): Aspekte einer Fachdidaktik Körperpflege. Ein Beitrag zur Standardentwicklung
ISBN 978-3-8325-5659-4 45.00 EUR

355 Livia Murer (2023): Diagnose experimenteller Kompetenzen beim praktisch-naturwissenschaftlichen Arbeiten. Vergleich verschiedener Methoden und kognitive Validierung eines Testverfahrens
ISBN 978-3-8325-5657-0 41.50 EUR (open access)

356 Andrea Maria Schmid (2023): Authentische Kontexte für MINT-Lernumgebungen. Eine zweiteilige Interventionsstudie in den Fachdidaktiken Physik und Technik
ISBN 978-3-8325-5605-1 57.00 EUR (open access)

357 Julia Ortmann (2023): Bedeutung und Förderung von Kompetenzen zum naturwissenschaftlichen Denken und Arbeiten in universitären Praktika
ISBN 978-3-8325-5670-9 37.00 EUR (open access)

358 Axel-Thilo Prokop (2023): Entwicklung eines Lehr-Lern-Labors zum Thema Radioaktivität. Eine didaktische Rekonstruktion
ISBN 978-3-8325-5671-6 49.50 EUR

359 Timo Hackemann (2023): Textverständlichkeit sprachlich variierter physikbezogener Sachtexte
ISBN 978-3-8325-5675-4 41.50 EUR (open access)

360 Dennis Dietz (2023): Vernetztes Lernen im fächerdifferenzierten und integrierten naturwissenschaftlichen Unterricht aufgezeigt am Basiskonzept Energie. Eine Studie zur Analyse der Wirksamkeit der Konzeption und Implementation eines schulinternen Curriculums für das Unterrichtsfach „Integrierte Naturwissenschaften 7/8“
ISBN 978-3-8325-5676-1 49.50 EUR

361 Ann-Katrin Krebs (2023): Vielfalt im Physikunterricht. Zur Wirkung von Lehrkräftefortbildungen unter Diversitätsaspekten
ISBN 978-3-8325-5672-3 65.50 EUR (open access)

362 Simon Kaulhausen (2023): Strukturelle Ursachen für Klausurmisserfolg in Allgemeiner Chemie an der Universität
ISBN 978-3-8325-5699-0 37.50 EUR (open access)

363 Julia Eckoldt (2023): Den (Sach-)Unterricht öffnen. Selbstkompetenzen und motivationale Orientierungen von Lehrkräften bei der Implementation einer Innovation untersucht am Beispiel des Freien Explorierens und Experimentierens
ISBN 978-3-8325-5663-1 48.50 EUR (open access)

364 Albert Teichrew (2023): Physikalische Modellbildung mit dynamischen Modellen
ISBN 978-3-8325-5710-2 58.50 EUR (open access)

365 Sascha Neff (2023): Transfer digitaler Innovationen in die Schulpraxis. Eine explorative Untersuchung zur Förderung der Implementation
ISBN 978-3-8325-5687-7 59.00 EUR (open access)

366 Rahel Schmid (2023): Verständnis von Nature of Science-Aspekten und Umgang mit Fehlern von Schüler*innen der Sekundarstufe I. Am Beispiel von digital-basierten Lernprozessen im informellen Lernsetting Smartfeld
ISBN 978-3-8325-5722-5 53.50 EUR (open access)

367 Dennis Kirstein (2023): Individuelle Bedingungs- und Risikofaktoren für erfolgreiche Lernprozesse mit kooperativen Experimentieraufgaben im Chemieunterricht. Eine Untersuchung zum Zusammenhang von Lernvoraussetzungen, Lerntätigkeiten, Schwierigkeiten und Lernerfolg beim Experimentieren in Kleingruppen der Sekundarstufe I
ISBN 978-3-8325-5729-4 50.50 EUR (open access)

368 Frauke Düwel (2024): Argumentationslinien in Lehr-Lernkontexten. Potenziale englischer Fachtexte zur Chromatografie und deren hochschuldidaktische Einbindung
ISBN 978-3-8325-5731-7 63.00 EUR (open access)

369 Fabien Güth (2023): Interessenbasierte Differenzierung mithilfe systematisch variierter Kontextaufgaben im Fach Chemie
ISBN 978-3-8325-5737-9 48.00 EUR (open access)

370 Oliver Grewe (2023): Förderung der professionellen Unterrichtswahrnehmung und Selbstwirksamkeitsüberzeugungen hinsichtlich sprachsensibler Maßnahmen im naturwissenschaftlichen Sachunterricht. Konzeption und Evaluation einer video- und praxisbasierten Lehrveranstaltung im Masterstudium
ISBN 978-3-8325-5738-6 44.50 EUR (open access)

371 Anna Nowak (2023): Untersuchung der Qualität von Selbstreflexionstexten zum Physikunterricht. Entwicklung des Reflexionsmodells REIZ
ISBN 978-3-8325-5739-3 59.00 EUR (open access)

372 Dominique Angela Holland (2023): Bildung für nachhaltige Entwicklung (BNE) kooperativ gestalten. Vergleich monodisziplinärer und interdisziplinärer Kooperation von Lehramtsstudierenden bei der Planung, Durchführung und Reflexion von Online-BNE-Unterricht
ISBN 978-3-8325-5760-7 47.00 EUR (open access)

373 Renan Marcello Vairo Nunes (2024): MINT-Personal an Schulen. Eine Untersuchung der Arbeitssituation und professionellen Kompetenzen von MINT-Lehrkräften verschiedener Ausbildungswege
ISBN 978-3-8325-5778-2 51.00 EUR (open access)

374 Mats Kieserling (2024): Digitalisierung im Chemieunterricht. Entwicklung und Evaluation einer experimentellen digitalen Lernumgebung mit universeller Zugänglichkeit
ISBN 978-3-8325-5786-7 45.50 EUR

375 Cem Aydin Salim (2024): Die Untersuchung adaptiver Lernsettings im Themenbereich „Schwimmen und Sinken“ im naturwissenschaftlichen Unterricht
ISBN 978-3-8325-5787-4 49.00 EUR (open access)

376 Novid Ghassemi (2024): Evaluation eines Lehramtsmasterstudiengangs mit dem Profil Quereinstieg im Fach Physik. Erkenntnisse zu Eingangsbedingungen, professionellen Kompetenzen und Aspekten individueller Angebotsnutzung
ISBN 978-3-8325-5789-8 41.50 EUR (open access)

377 Martina Flurina Cavelti (2024): Entwicklung und Validierung eines Messinstruments zur Erfassung der Schülerkompetenzen im Bereich des wissenschaftlichen Skizzierens im Fach Chemie in der Sekundarstufe I
ISBN 978-3-8325-5829-1 45.00 EUR (open access)

378 Tom Bleckmann (2024): Formatives Assessment auf Basis von maschinellem Lernen. Eine Studie über automatisiertes Feedback zu Concept Maps aus dem Bereich Mechanik
ISBN 978-3-8325-5842-0 46.50 EUR (open access)

379 Jana Marlies Rehberg (2024): Das physikspezifische Mindset zum Studienbeginn. Fragebogenentwicklung und Aufbau einer Online-Intervention
ISBN 978-3-8325-5850-5 59.50 EUR (open access)

380 Florian Trauten (2024): Entwicklung und Evaluation von automatisierten Feedbackschleifen in Online-Aufgaben im Fach Chemie
ISBN 978-3-8325-5859-8 46.00 EUR (open access)

381 Johanna Dejanovikj (2024): Reflexion als Lerngegenstand im Unterricht zur Förderung von Bewertungskompetenz
ISBN 978-3-8325-5860-4 41.00 EUR (open access)

382 Katharina Flieser (2024): Verständlichkeit physikalischer Sachtexte. Untersuchungen zum Wirkungsgefüge zwischen sprachlicher Textgestaltung und der Behaltensleistung sowie der Textwahrnehmung im Schulfach Physik
ISBN 978-3-8325-5858-1 44.00 EUR (open access)

383 Stephanie Neppl (2024): Perspektivenübernahme im Physikunterricht. Explorative Interviewstudie zu einer Seminarkonzeption mit dem Schwerpunkt Perspektivenübernahme bei der Planung von Physikunterricht
ISBN 978-3-8325-5865-9 48.00 EUR (open access)

384 Katja Plicht (2024): Ein Physikübungskonzept zur Förderung der Problemlösekompetenz. Entwicklung und empirische Evaluation eines Strategietrainings auf der Basis von Expertisemerkmalen
ISBN 978-3-8325-5875-8 45.00 EUR (open access)

385 Svenja Boegel (2024): Feedback beim Experimentieren: Zur Rolle von Cognitive Load und Motivation
ISBN 978-3-8325-5911-3 45.50 EUR (open access)

386 Laura Pannullo (2025): Wahlmöglichkeiten beim Experimentieren: Entwicklung und Erprobung eines Konzeptes für Experimente in inklusiven Lerngruppen im Physikunterricht
ISBN 978-3-8325-5916-8 46.50 EUR (open access)

387 Matthias Schweinberger (2025): „Stumme Videos“. Aufmerksamkeitslenkende Moderation von Demonstrationsexperimenten
ISBN 978-3-8325-5930-4 50.50 EUR

388 Rike Große-Heilmann (2025): Entwicklung fachdidaktischen Wissens zum Einsatz digitaler Medien im Fach Physik
ISBN 978-3-8325-5943-4 51.00 EUR (open access)

389 Lion Cornelius Glatz (2025): Eine vergleichende Betrachtung der Überzeugungskraft von Experimenten zum Teilchenmodell
ISBN 978-3-8325-5954-0 42.00 EUR (open access)

390 Rasmus Viefers (2025): Förderung der Variablenkontrollstrategie im Sachunterricht
ISBN 978-3-8325-5968-7 44.50 EUR (open access)

391 Luc Broder Albrecht (2025): Kritisches Denken in der Chemielehrer*innenbildung. Eine konzeptionell-empirische Studie zur Integration und Operationalisierung
ISBN 978-3-8325-5970-0 57.00 EUR (open access)

392 Anna Weißbach (2025): Reflexion von Physikunterricht. Entwicklung und Validierung eines Testinstruments mit Assessment-Feedback zur Reflexionsfähigkeit Physik-Lehramtsstudierender
ISBN 978-3-8325-5977-9 51.50 EUR (open access)

393 Robert Gieske (2025): Förderung fachwissenschaftlicher und kommunikativer Kompetenzen von Schüler*innen durch Disaggregate Instruction im Fach Chemie. Konzeption und Durchführung einer vergleichenden Interventionsstudie zur Analyse der Wirksamkeit sprachsensibler Unterrichtsansätze in der Sekundarstufe I
ISBN 978-3-8325-5984-7 51.00 EUR (open access)

394 Thomas Sean Weatherby (2025): Talking Circuits. The Development and Assessment of a Digitally-Scaffolded, Collaborative Method for Teaching and Learning Electrical Circuits in Early Secondary Schools
ISBN 978-3-8325-6008-9 57.00 EUR (open access)

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

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

Kritisches Denken ist eine zentrale Voraussetzung, um naturwissenschaftliches Wissen in gesellschaftlichen Entscheidungssituationen handlungswirksam zu nutzen. Trotz seiner prominenten Stellung in Bildungsdiskursen bleibt es diffus definiert und schwer adressierbar.

Die vorliegende Arbeit entwickelt auf konzeptioneller Ebene eine interdisziplinär fundierte Definition Kritischen Denkens, die kognitive und affektive Dimensionen integriert, und führt zwei strukturierende Modelle ein. Darauf aufbauend wird empirisch untersucht, wie sich Kritisches Denken erschließen und operationalisieren lässt. In einem iterativen Forschungsprozess wurde ein triangulierendes Instrument zur Analyse multiperspektivischer Bewertungsprozesse entwickelt und erprobt. Erfasst werden fachliche, epistemologische und selbstreflexive Aspekte Kritischen Denkens bei der Bewertung gesellschaftlicher Problemstellungen mit naturwissenschaftlichem Bezug. Die Erhebungen erstreckten sich über vier Semester und umfassen 97 Studierende mit 143 Fragebögen und 51 Interviews.

Die Ergebnisse zeigen, dass Studienfortschritt allein nicht zur Erweiterung des Denkens führt. Erst durch die explizite Thematisierung Kritischen Denkens lassen sich signifikante Veränderungen im Reflexionsverhalten und in der Perspektivenvielfalt feststellen. Dies verdeutlicht die Notwendigkeit einer gezielten Auseinandersetzung mit dem Thema und eröffnet sowohl für die Forschung neue Perspektiven zur Analyse von Denkprozessen als auch für die Lehrerkräftebildung konkrete Ansatzpunkte zur Förderung Kritischen Denkens.

Logos Verlag Berlin

ISBN 978-3-8325-5970-0