

Katharina Muth



Bewertungskriterien ethischer und religiöser Urteilskompetenz

Eine qualitative Studie über Prüfungsaufgaben
und Bewertungsvorgaben im schriftlichen Abitur
des Faches Evangelische Religionslehre



Bewertungskriterien ethischer und religiöser Urteilskompetenz

STUDIEN ZUR RELIGIÖSEN BILDUNG (StRB)

Herausgegeben von
Michael Wermke und Thomas Heller

Band 22

Katharina Muth

Bewertungskriterien ethischer und religiöser Urteilskompetenz

Eine qualitative Studie über Prüfungsaufgaben
und Bewertungsvorgaben im schriftlichen Abitur
des Faches Evangelische Religionslehre



EVANGELISCHE VERLAGSANSTALT
Leipzig



Katharina Muth, Jg. 1991, studierte Deutsch und Evangelische Religionslehre für Gymnasium an der Friedrich-Schiller-Universität Jena. Sie ist Mitglied in der Gesellschaft für wissenschaftliche Religionspädagogik (GwR) und des Zentrums für Religionspädagogische Bildungsforschung (ZRB). 2021 wurde sie mit der vorliegenden Arbeit promoviert.

Bibliographische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

© 2021 by Evangelische Verlagsanstalt GmbH · Leipzig
Printed in Germany

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Das Buch wurde auf alterungsbeständigem Papier gedruckt.

Cover: Kai-Michael Gustmann, Leipzig
Satz: 3w+p, Rimpf
Druck und Binden: Hubert & Co., Göttingen

ISBN Print 978-3-374-06936-1 // eISBN (PDF) 978-3-374-06937-8
www.eva-leipzig.de

Vorwort

Als jemand, der 2010 sein Studium für Lehramt begonnen hat, wurden mir Bildungsmonitoring, Standardisierung und Kompetenzorientierung als wesentliche Elemente der Qualitätsentwicklung für Schule vermittelt, welchen es zu folgen gelte. Ich erinnere mich an eine Vorlesung der Erziehungswissenschaft, in der der Professor scharf äußerte, dass ein operierender Arzt auch nicht mit alten Methoden der Blinddarmentnahme gewohnheitsgemäß fortfahren könne, sobald es empirische Belege dafür gäbe, dass eine neuere Methode effizienter sei und sicherer zum Ziel führt. Lehrerinnen und Lehrer, so sein Fazit, sind ebenso verpflichtet, empirisch gesicherte Kompetenzmodelle ernst zu nehmen und routinierte Methoden und Ziele zu hinterfragen und gegebenenfalls zu ändern. Dies schien mir evident und so nahm ich mir vor, gemäß dem Gelernten und der bildungsadministrativen Vorgaben den Lernprozess vom Learning-Outcome her zu denken, genau zu diagnostizieren, auf welcher Ebene der Kompetenzentwicklung sich die Schülerinnen und Schüler befinden und Lernimpulse zu schaffen, durch welche sie auf die nächste Niveaustufe ihrer Kompetenz gelangen können. Am Ende meines Studiums war mir das Anliegen der Kompetenzorientierung klar, aber anhand welcher Qualitätskriterien man die Entwicklung einer konkreten von den bildungsadministrativen Vorgaben zwar bestimmten, aber nicht näher definierten Kompetenz nun erheben kann, blieb mir offen. Die vorliegende Studie ist als eine für mein Selbstverständnis als angehende Lehrkraft bedeutsame Frage nach der Umsetzbarkeit der auf bildungsadministrativer Ebene erhobenen Anforderungen an Qualitätssicherung und diagnostische Validität zu verstehen. Sie stellt eine Auseinandersetzung mit dem Paradigma der Kompetenzorientierung dar, welches – anders als empirisch bestätigte Methoden der Blinddarmentnahme – durchaus im Forschungsdiskurs auch mit grundlegender Kritik konfrontiert ist. Unbenommen von dieser Kritik ist mein – in der Kompetenzorientierung verankerte Anspruch – nach einem Religionsunterricht zu fragen, der die Schülerinnen und Schüler herausfordert, ihre Kompetenzen weiter auszubauen und der benennen kann, welche Schritte zu gehen sind, um Kompetenzaneignung zu fördern und valide zu diagnostizieren. Die Studie fragt daher exemplarisch nach den Qualitätsmerkmalen einer meines Erachtens sehr wesentlichen und komplexen Kompetenz – nämlich der Urteilskompetenz. Ur-

teilskompetenz trägt maßgeblich zum übergeordneten Ziel der Schule bei: zur Mündigkeit der Lernenden, die gesellschaftliche und soziale Partizipation ermöglicht. Sie spielt daher in allen Fächern – wenn auch mit unterschiedlicher Ausprägung – eine Rolle, so dass die Frage nach Qualitätsmerkmalen und Bewertungskriterien von Urteilsbildung nicht auf den Religionsunterricht begrenzt, sondern von fachübergreifendem Interesse ist.

Die Studie fragt in erster Linie nach den Strukturvorstellungen von Urteilsbildung, die schriftlichen Abituraufgaben zugrunde liegen, nach Standards, nach der Bedeutung der Bekenntnisorientierung des Faches Religion im Urteilsbildungsprozess und nach diagnostischer Güte. Dieses Wissen über die Beschaffenheit der anzustrebenden Kompetenz ist meines Erachtens für den Religionsunterricht eine notwendige, jedoch keine hinreichende Voraussetzung für dessen Qualität. Diese muss sich neben dem Erreichen von Niveaustufen auch an dem Erreichen einer resonanzerzeugenden gemeinsamen Suchbewegung zwischen Lehrerinnen und Lehrern, Schülerinnen und Schülern und dem Inhalt messen lassen. Ich bin überzeugt, dass sich diese beiden Arten, Unterrichtsqualität zu messen, nicht ausschließen, sondern in einen fruchtbaren Einklang zu bringen sind.

Das Vorwort abschließend sei noch ein Dank ausgesprochen an das THÜRINGER MINISTERIUM FÜR UNTERRICHT, JUGEND UND SPORT, das BAYERISCHE STAATSMINISTERIUM FÜR BILDUNG UND KULTUS sowie das NIEDERSÄCHSISCHE KULTUSMINISTERIUM. Sie haben das Forschungsprojekt durch die Bereitstellung des Untersuchungsmaterials in Form von Abituraufgabenstellungen und der dazugehörigen Erwartungshorizonte sowie durch die Zusammenstellung statistischer Daten unterstützt und ermöglicht. Die Publikation der Studie inklusive der Open-Access-Veröffentlichung wurde durch eine großzügige finanzielle Unterstützung der Doktorandenschule des Forschungszentrums Laboratorium Aufklärung sowie der Evangelischen Kirche in Mitteldeutschland möglich, wofür ich mich recht herzlich bedanken möchte.*

Darüber hinaus bin ich für die anregende Unterstützung und die Begleitung auf dem Entstehungsprozess der Studie meinem Doktorvater Prof. Dr. Michael Wermke, meiner Zweitgutachterin Prof. Dr. Bärbel Kracke, dem Team am Lehrstuhl für Religionspädagogik der Friedrich-Schiller-Universität Jena, insbesondere PD Dr. Thomas Heller, Dr. Sophie Seher und AR Dr. Marita Koerrenz, sowie den zahlreichen Korrekturlesenden, vor allem Michael und Eduard Spinka, Damaris und Daniel Muth, sehr dankbar. Meiner Familie und vor allem meinem Mann Willi Kaufmann danke ich für Zuspruch und Rückhalt.

Herzlichen Dank.
Katharina Muth

Jena, im Februar 2021

* Im Internet ist die Studie unter folgender URL abrufbar: https://www.eva-leipzig.de/images/content/Muth_StRB-22.pdf

Inhalt

Vorwort	5
1 Einleitung	13
2 Urteilskompetenz als komplexe Anforderung – Erkenntnisinteresse der Studie	19
2.1 Definition des Untersuchungsgegenstandes	19
2.2 Verortung der Studie in einem kompetenzorientierten Bildungsverständnis	29
2.2.1 Zentrale Aspekte eines kompetenzorientierten Bildungsverständnisses	29
2.2.2 Der Konstruktcharakter von Kompetenzen	35
2.3 Verortung der Studie im religionspädagogischen Diskurs	38
2.4 Forschungsanliegen und Forschungsfragen	48
3 Diagnostische Rahmenbedingungen von Urteilskompetenz in schriftlichen Abiturprüfungen – Forschungshintergrund	49
3.1 Vergleichbarkeit zwischen zentralen Abschlussprüfungen	49
3.2 Vorgaben der Einheitlichen Prüfungsanforderungen für das Fach Evangelische Religionslehre	54
3.2.1 Urteilsfähigkeit im Kompetenzmodell der Einheitlichen Prüfungsanforderungen	55
3.2.2 Urteilsfähigkeit innerhalb der Anforderungsbereiche ..	57
3.2.3 Exkurs: Die Perspektive der EKD – Urteilsfähigkeit zur Orientierungsfindung	60

3.3 Diagnostische Grundlagen und Besonderheiten zur Urteilsfähigkeit im Kontext von Abschlussprüfungen	62
3.3.1 Formen von Kompetenzmodellen	62
3.3.2 Bezugsnormen, Gütekriterien und Bewertungsansätze	64
3.3.3 Urteilsbildung im Medium der Schriftlichkeit	70
4 Qualitätsmerkmale ethischer und religiöser Urteilskompetenz – Aktueller Forschungsstand	73
4.1 Allgemeinpädagogische Forschungen	74
4.1.1 Argumentationstheorie	74
4.1.2 Moralische Urteilsbildung – Lawrence Kohlberg und Georg Lind	77
4.1.3 Teilkompetenzen ethisch-moralischer Kompetenzniveaus – Das Projekt ETiK	80
4.1.4 Forschungen zu Bewertungskompetenz aus der Biologiedidaktik	83
4.2 Religionspädagogischer Forschungsstand	87
4.2.1 Religiös-weltanschauliche Neutralität des Staates in Urteilsfragen? Eine rechtliche Perspektive	87
4.2.2 Modellbildungen ethischer Urteilsbildung	91
4.2.3 Empirische Studien zur Argumentations- und Urteilsfähigkeit von Schülerinnen und Schülern	94
4.3 Zusammenfassung	99
5 Forschungsdesign	103
5.1 Material	103
5.2 Methode	109
5.2.1 Definition der Auswahl- und Analyseeinheit sowie des Aufgabenbegriffs	111
5.2.2 Auswahl der Aufgabenstellungen zum Kompetenzbereich ethische und religiöse Urteilsfähigkeit	114
5.2.3 Grenz- und Sonderfälle	115

5.3 Kategoriensystem der Studie	123
5.3.1 Quantitative Aspekte	125
5.3.2 Kategorien zur Analyse der Aufgabenstellungen	125
5.3.3 Kategorien zur Analyse der Erwartungshorizonte	135
5.4 Analysevorgehen und Methodenreflexion	140
5.4.1 Beantragung der Einsicht in die Abituraufgabenstellungen	141
5.4.2 Auswahl der Aufgabenstellungen zur Urteilsbildung ...	141
5.4.3 Erstellung des Kategoriensystems und Auswertung der Daten	142
6 Kategorienorientierte Ergebnisse	147
6.1 Ergebnisse in Bezug auf die quantitativen Aspekte	147
6.1.1 Die Gesamtaufgabenzahl	148
6.1.2 Prozentualer Anteil ethischer Urteilsbildung	149
6.2 Ergebnisse in Bezug auf die Aufgabenstellungen	150
6.2.1 Aufgabenart (A)	150
6.2.2 Fragestellungen der Urteilsbildung (B)	151
6.2.3 Reflexionskriterien (C)	159
6.2.4 Exkurs: Rolle christlicher Sichtweisen in den Aufgabenstellungen zur Urteilsfähigkeit	161
6.2.5 Textsorten der Materialien, die den Aufgaben zur Urteilsbildung zugrunde liegen (D)	162
6.3 Ergebnisse in Bezug auf die Erwartungshorizonte	164
6.3.1 Passung zwischen geforderter und erwarteter Urteilsfähigkeit (E)	164
6.3.2 Qualitätsmerkmale von Urteilsbildung in den Erwartungshorizonten (F)	169
6.3.3 Explizite Hinweise zum Umgang mit Schülermeinungen (G)	175
6.3.4 Formulierung konträrer Positionen in den Erwartungshorizonten (H)	176
6.4 Zusammenfassung der kategorienorientierten Ergebnisse	176

7	Fallorientierte Ergebnisse	179
7.1	Methodische Einführung	179
7.2	Formale Ausgestaltung der Erwartungshorizonte im Vergleich zwischen den untersuchten Bundesländern	181
7.2.1	Formale Aspekte bayerischer Erwartungshorizonte	181
7.2.2	Formale Aspekte niedersächsischer Erwartungshorizonte	181
7.2.3	Formale Aspekte thüringischer Erwartungshorizonte	183
7.2.4	Zusammenfassung	184
7.3	Fallanalysen aus Bayern	184
7.3.1	Die Frage nach dem Maßstab	184
7.3.2	Die Frage nach Bewertungskriterien	186
7.3.3	Die Frage nach der Ausformulierung von Pro- und Kontrapositionen	188
7.4	Fallanalysen aus Niedersachsen	190
7.4.1	Ausformulierung von Pro- und Kontrapositionen	190
7.4.3	Urteilsbildung zu gegenwartsbezogenen gesellschaftspolitischen Fragestellungen	193
7.4.4	Exkurs: Fehlende Schwierigkeitsdifferenzierung der Anforderungsniveaus im Bereich Urteilsfähigkeit	195
7.5	Fallanalysen aus Thüringen	198
7.5.1	Komplexität und Offenheit der Aufgabenstellung	198
7.5.2	Exkurs: Unstimmigkeit zwischen Operatoren und Erwartungen	202
7.5.3	Urteilskompetenz im Zusammenhang mit einer Gedichtinterpretation	204
7.6	Zusammenfassung der fallorientierten Ergebnisse	207
8	Aufgabenkulturen und Bewertungsvorgaben – Profilierung der Bundesländer	209
8.1	Bayern – Christliche Weltsicht als Reflexionsmedium ethischer Fragestellungen	209
8.1.1	Aufgabenstellungen zu Urteilsfähigkeit aus Bayern	209
8.1.2	Erwartungshorizonte zu Urteilsaufgaben aus Bayern	210

8.2	Niedersachsen – Durch Theologinnen und Theologen vermittelte Weltsicht in formalisiertem Rahmen	211
8.2.1	Aufgabenstellungen zu Urteilsfähigkeit aus Niedersachsen	211
8.2.2	Erwartungshorizonte zu Urteilsaufgaben aus Niedersachsen	213
8.3	Thüringen – Komplexe Aufgabenstellungen unter Einbezug biblischer Texte	214
8.3.1	Aufgabenstellungen zu Urteilsfähigkeit aus Thüringen	214
8.3.2	Erwartungshorizonte zu Urteilsbildungsaufgaben aus Thüringen	215
9	Einheit in Vielfalt? Eine Diskussion der Ergebnisse im Spiegel des Anspruchs auf Vergleichbarkeit	219
9.1	Aspekte der Prüfungsgestaltung – Kontextbedingte Differenzen vs. kontextunabhängige Differenzen	220
9.1.1	Differenzen als Chance einer kontextsensiblen Prüfung	221
9.1.2	Differenzen als Grenzen für eine gerechte Leistungsbewertung	224
9.1.3	Die Kunst der Unterscheidung	228
9.2	Aspekte der Zielvorgaben – Gütekriterien und Bewertungsmaßstäbe für Urteilsfähigkeit in den Erwartungshorizonten	229
9.2.1	Qualitätsmerkmale der Erwartungshorizonte – Differenziertheit, Eigenständigkeit, fundierte Sachkenntnisse, Schlüssigkeit und sprachliche Gestaltung	230
9.2.2	Gütekriterien und Bewertungsmaßstäbe in der Forschungsliteratur – Multiperspektivität, Transparenz der zugrunde liegenden Werte, Kontexteinbettung, Adressatenbezug, Folgenreflexion	232
9.2.3	Die Erwartungshorizonte vor dem Hintergrund von Messtheorien	237
9.3	Bekenntnisorientierung im Kontext Urteilsbildung	242
9.3.1	Bekenntnisorientierung im Kontext theologischer Fragestellungen	243

9.3.2 Bekenntnisorientierung in Bezug auf Reflexionskriterien	244
10 Résumé und Ausblick	249
Literaturverzeichnis	253
Abbildungsverzeichnis	267
Tabellenverzeichnis	269
Abkürzungsverzeichnis	271
Anhang	273
Anhang I: Berücksichtigte Operatoren aus den EPA	273
Anhang II: Die von den EPA vorgegebenen Bewertungskriterien	274
Anhang III: Die Fälle der Studie (42) – Aufgabenstellungen zu Urteilsfähigkeit in kontroversen ethischen oder religiösen Fragestellungen aus der Auswahleinheit	275
Anhang IV: Urteilsbildungen mit Perspektivübernahme	280
Anhang V: Gedicht »Stufen« von Hermann Hesse	282

1 Einleitung

Die Ausrichtung des Religionsunterrichts auf gesellschaftsrelevante Fragestellungen und das Ziel, Schülerinnen und Schüler zu (religiöser) Mündigkeit zu befähigen, ist seit der Entwicklung des Problemorientierten Religionsunterrichts in den 1960er konzeptionell tief im Selbstverständnis des Religionsunterrichts verankert. So erhebt der Religionsunterricht den Anspruch, »die Auseinandersetzung mit Ideologien, Weltanschauungen und Religionen [zu] ermöglichen«, »für die Verständigung über gesellschaftliche Grundorientierungen unverzichtbar« zu sein und die Schülerinnen und Schüler dazu zu befähigen, »in religiösen und ethischen Fragen begründet urteilen« zu können.¹ Zur Erreichung dieses Anspruchs bedarf es der ethischen und religiösen Urteilsfähigkeit, welche von den Einheitlichen Prüfungsanforderungen neben Wahrnehmungs- und Darstellungsfähigkeit, Deutungsfähigkeit, Dialogfähigkeit und Gestaltungsfähigkeit als eine von fünf zentralen Kompetenzen benannt wird, die im Religionsunterricht erworben werden sollen.² Doch was macht diese Kompetenz aus? Wie kann sie

¹ Vgl. STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER (Hrsg.), Einheitliche Prüfungsanforderungen in der Abiturprüfung Evangelische Religionslehre. Beschluss vom 01.12.1989 i. d. F. vom 16.11.2006. Abrufbar unter: https://www.kmk.org/fileadmin/veroeffentlichungen_beschlue-sse/1989/1989_12_01-EPA-Ev-Religion.pdf (13.04.2019), 5 sowie 8.

² Vgl. a. a. O., 8 f.

Ebenso legt die EVANGELISCHE KIRCHE IN DEUTSCHLAND in ihren Texten zum Religionsunterricht das von den EPA vertretene Kompetenzmodell zugrunde, so dass das fünfgliedrige Modell auch von den Curricula der meisten Bundesländer aufgegriffen wurde (vgl. EVANGELISCHE KIRCHE IN DEUTSCHLAND (Hrsg.), Kompetenzen und Standards für den Evangelischen Religionsunterricht der Sekundarstufe I. Ein Orientierungsrahmen. Texte 111. Hannover: 2011. Sowie DIES. (Hrsg.), Kerncurriculum für das Fach Evangelische Religionslehre in der gymnasialen Oberstufe. Themen und Inhalte für die Entwicklung von Kompetenzen religiöser Bildung. Texte 109. Hannover 2010).

gefördert werden, um dem oben formulierten Anspruch tatsächlich gerecht zu werden? Und: Welche Qualitätsmerkmale vermögen es, einen Kompetenzzuwachs im Bereich Urteilsfähigkeit messbar und bewertbar zu machen? Diese Frageperspektive, welche die Beschreibung von Qualitätsmerkmalen und graduell fassbaren Niveaustufen akzentuiert, basiert auf einem kompetenzorientierten Bildungsverständnis. Dieses wird in der deutschsprachigen Schulpädagogik bereits seit dem Erscheinen der sogenannten Klieme-Expertise 2003 intensiv diskutiert.³ Auch für den Religionsunterricht hat diese kontrovers geführte Diskussion einen weitreichenden Impuls im Nachdenken darüber mit sich gebracht, welchen Ertrag – sogenannte Outcomes – der Religionsunterricht leisten kann und ob und wenn ja, wie sich die Kompetenzen beschreiben lassen, die im Religionsunterricht erworben werden sollen (vgl. Kapitel 2.2 und 2.3). Eine differenzierte Vorstellung der zu fördernden Kompetenzen ist vor allem für die Leistungsbewertung zentral: Um prüfen und bewerten zu können, in welchem Maße jemand über Urteilsfähigkeit verfügt, muss man wissen, was Urteilsfähigkeit ausmacht. Schulische Leistungsmessung wurde in der Religionspädagogik zwar bereits Ende der 1970er Jahre breit diskutiert. Zwei Standardwerke, die die Diskussion wesentlich beeinflusst haben, sind die 1979 erschienen Schriften *Leistungsmessung im Religionsunterricht. Methoden und Beispiele* von BERNHARD JENDORFF sowie *Religionsunterricht in der Leistungsschule. Gutachten, Dokumente* von KARL ERNST NIPKOW.⁴ Jedoch bezog sich dieser bereits vor 40 Jahren angestoßenen Diskurs über Leistungsmessung im Religionsunterricht im Wesentlichen auf Fragen nach dem Umgang mit Leistungsdruck und dem Zwiespalt zwischen Bewertung und der Rechtfertigungslehre, die eine leistungsunabhängige, unbedingte Rechtfertigung vor Gott betont. OLIVER REIS und THERESA SCHWARZKOPF konstatieren zum religionspädagogischen Diskurs:

Da die zitierten *Kompetenzmodelle* von *Urteilsfähigkeit* sprechen, werden auch in der vorliegenden Studie die Begriffe *Urteilsfähigkeit* und *Urteilskompetenz* synonym gebraucht. Zum Kompetenzbegriff, der sich auf komplexe Fähigkeiten bezieht vgl. Kapitel 2.2.

³ Die letzte Reform des deutschen Bildungssystems wurde durch die Expertise »Zur Entwicklung nationaler Bildungsstandards« festgeschrieben, welche vom Bundesministerium für Bildung und Forschung in Auftrag gegeben wurde. Federführend bei der Entwicklung dieser Expertise war ECKHARD KLIEME, wodurch das Dokument als die »Klieme-Expertise« in den Diskurs eingegangen ist. Diese Bezeichnung soll im Folgenden übernommen werden. Vgl. BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (BMBF) (Hrsg.), *Zur Entwicklung nationaler Bildungsstandards. Expertise*. Bonn/Berlin 2007, 72.

⁴ BERNHARD JENDORFF, *Leistungsmessung im Religionsunterricht. Methoden und Beispiele*. München 1979. KARL ERNST NIPKOW, *Religionsunterricht in der Leistungsschule. Gutachten, Dokumente*. Gütersloh 1979.

»Im Vordergrund stehen [...] noch nicht die Diagnose der Lernausgangslage oder des Lernerfolgs, sondern die Bedingungen der religionspädagogischen Angemessenheit der Leistungsmessung. Dieser religionspädagogische Vorbehalt führt zur Betonung der Feedbackkultur, der Grenzen der Messbarkeit und der Ablehnung der Rede vom Fehler.«⁵

In diesem Zusammenhang kritisieren REIS und SCHWARZKOPF, dass schwierigkeitskonstituierende Merkmale, Niveaustufendifferenzierungen, Qualitätsmarker und auch das, was man auf der Grundlage von Fakten eindeutig als falsch benennen kann, im religionspädagogischen Diskurs zu kurz kommen.⁶ Sowohl im Hinblick auf die Qualitätssicherung des Religionsunterrichts im Allgemeinen als auch im Kontext schriftlicher Abiturprüfungsaufgaben, die den Anspruch einer bundesweiten Vergleichbarkeit erheben, sind jedoch ein intensiver Diskurs und eine ausgeprägte Sprachfähigkeit über Qualitätsmerkmale als Bewertungsgrundlage von zentraler Bedeutung.

Die vorliegende Studie möchte einen Anstoß dafür geben, die Kompetenz Urteilsfähigkeit intensiver zu beforschen, indem sie auf der Grundlage aktueller Erkenntnisse über Urteilsfähigkeit aus (Pädagogischer) Psychologie und Didaktikforschung nach an Schülerinnen und Schüler gestellte Anforderungen im Zusammenhang mit Urteilsfähigkeit sowie nach deren Bewertungskriterien fragt.

Als Untersuchungsgegenstand dienen dafür schriftliche Abituraufgabenstellungen und ihre Erwartungshorizonte der exemplarisch ausgewählten Bundesländer Bayern, Niedersachsen und Thüringen. Abituraufgabenstellungen und ihre Erwartungshorizonte sind im Zusammenhang mit der Frage nach Kompetenzerwartungen und Bewertungskriterien besonders interessant, da sie Auskunft über die Standards geben, die von Schülerinnen und Schülern erwartet werden, nachdem ihre schulische Bildung als abgeschlossen gilt. Sie markieren somit einen Maximalanspruch dessen, was erwartet wird. Damit rückt die Studie ein Material in den Fokus, welches in der Religionspädagogik bisher noch unerforscht ist und von dem zu erwarten ist, dass es auch Einfluss auf die Kompetenzerwartungen und Zielvorstellungen im Unterricht der Sekundarstufe II nimmt. VARVARA DISDORN-LIESEN schreibt in Bezug auf Abiturprüfungen im Fach Deutsch: »Die Aufgabenkonstruktionen, bestehend aus Prüfungsaufgaben und Bewertungsvorgaben, nehmen in erheblichem Maße Einfluss auf die Unter-

⁵ OLIVER REIS/THERESA SCHWARZKOPF, Diagnosemodelle für den Religionsunterricht. Eine Bestandsaufnahme, in: Religionspädagogische Beiträge 76 (2017), 85–95, hier 86.

⁶ Vgl. OLIVER REIS/THERESA SCHWARZKOPF, Diagnose religiöser Lernprozesse. Ein kompetenzdiagnostisches Grundlagenmodell, in: DIES. (Hrsg.), Diagnose im Religionsunterricht. Konzeptionelle Grundlagen und Praxiserprobungen. Berlin 2015, 15–120. Vgl. insbesondere 19, 37 f. sowie 42–45.

richtsgestaltung; sie sind Bezugspunkte für die unterrichtspraktische Arbeit und werden von Lehrkräften, Schülerinnen und Schülern sowie Eltern als Orientierungsrahmen angesehen.«⁷ Für den Religionsunterricht ist der Einfluss der schriftlichen Abiturprüfungen auf die Unterrichtspraxis sicherlich deshalb zu relativieren, da eine schriftliche Abiturprüfung im Fach Evangelische Religionslehre – im Gegensatz zum Fach Deutsch – in der Regel fakultativ ist und an manchen Schulen gar nicht angeboten wird (vgl. Kapitel 5.1). Dennoch gibt die Analyse der Prüfpraxis einen sowohl interessanten als auch für weitere Forschungen anknüpfungsfähigen Einblick in Prüfungskonventionen und Bewertungsvorgaben religiöser und ethischer Urteilsfähigkeit, der auch Fragen hinsichtlich der unterrichtspraktischen Arbeit aufwirft und neu reflektieren lässt.

Der Aufbau der Studie folgt dabei der klassischen Darstellungsstruktur empirischer Forschungen: In Kapitel 2 wird das Erkenntnisinteresses der Studie entwickelt. Dazu wird zunächst der Forschungsgegenstand definiert und eingegrenzt (Kap. 2.1) sowie die Studie in einem kompetenzorientierten Bildungsverständnis und in der religionspädagogischen Debatte verortet (Kap. 2.2 und 2.3). Vor dem Hintergrund der Definitionen und Einordnungen werden – das Kapitel abschließend – die Forschungsfragen der Untersuchung formuliert (Kap. 2.4). Sie orientieren sich an zwei übergeordneten, aufeinander aufbauenden Fragestellungen: *Zum einen wird gefragt, mittels welcher Aufgabenstellungen und welcher Bewertungskriterien Urteilsbildung im schriftlichen Abitur des Faches Evangelische Religionslehre geprüft und bewertet wird. Zum anderen steht die Frage im Fokus, wie diese Prüfungskonventionen vor dem Hintergrund der diagnostischen Rahmenbedingungen von Abituraufgabenstellungen und des aktuellen Forschungsdiskurses über Urteilskompetenz zu bewerten sind.*

Zentrale Grundlagen zur Beantwortung dieser Fragestellungen sind die diagnostischen Rahmenbedingungen, welche die Prüfungskonventionen bestimmen und in Kapitel 3 als Forschungshintergrund Darstellung finden. Das Kapitel arbeitet sich von einer allgemeinen Betrachtung der Abiturprüfungsrahmenbedingungen in Kapitel 3.1 zum Speziellen: Kapitel 3.2 analysiert die Einheitlichen Prüfungsanforderungen für das Fach Evangelische Religionslehre hinsichtlich ihrer Vorgaben zur Prüfung von Urteilskompetenz, während Kapitel 3.3 die diagnostischen Besonderheiten von Urteilsfähigkeit reflektiert, welche sich durch den Kontext schriftlicher Abiturprüfungen ergeben. Zur Einordnung der Ergebnisse der Studie ist die Kenntnis über den aktuellen Forschungsstand zu Qualitätsmerkmalen von Urteilsbildung wesentlich. Dieser wird unter allgemeinpädagogischer Perspektive in Kapitel 4.1 und unter religionspädagogischer Perspektive in Kapitel 4.2 dargestellt.

⁷ Vgl. VARVARA DISDORN-LIESEN, Vergleichbarkeit in der Vielfalt. Leistungsanforderungen und Leistungsfeststellungen im Zentralabitur Deutsch. Wiesbaden 2016, 7 f.

In Kapitel 5 zum Forschungsdesign wird ausgehend vom untersuchten Material (Kap. 5.1) die Methodik der Untersuchung vorgestellt, welche sich an einer qualitativen Inhaltsanalyse nach UDO KUCKARTZ und PHILIPP MAYRING orientiert (vgl. Kapitel 5.2). Um aus den im Untersuchungszeitraum gestellten Abituraufgabenstellungen jene auszuwählen, die eine Urteilsbildung erfordern, wurde auf Grundlage von Kapitel 2 eine Definition erstellt, welche in Kapitel 5.2 dargelegt und diskutiert wird. Kapitel 5.3 dient der Darstellung des Kategoriensystems, welches in seinen deduktiv erstellten Hauptkategorien auf Darstellungen des Forschungshintergrundes und des aktuellen Forschungsstandes zurückgreift. Die Hauptkategorien wurden während der Analyse induktiv ergänzt.

Die Erhebung mittels des Kategoriensystems wurde auf zwei verschiedenen Wegen ausgewertet: Zum einen fanden die erstellten Kategorien einzeln Betrachtung. Eine Darstellung dieser kategorienorientierten Ergebnisse findet in Kapitel 6 statt, wobei Kapitel 6.1 quantitative Aspekte fokussiert, während Kapitel 6.2 Ergebnisse in Bezug auf die Aufgabenstellungen präsentiert und Kapitel 6.3. Ergebnisse in Bezug auf die Erwartungshorizonte verdeutlicht. In Kapitel 6.4 werden die kategorienorientierten Ergebnisse zusammengefasst.

Um die Kontextualisierung einzelner Aufgaben und Zusammenhänge zwischen den Kategorien besser darstellen zu können, schließt sich an die kategorienorientierte Auswertung zum anderen eine fallorientierte Analyse an, deren Ergebnisse in Kapitel 7 dargestellt werden. Durch die Betrachtung mehrerer Kategorien anhand exemplarisch ausgewählter Fälle werden die Ergebnisse aus Kapitel 6 in Kapitel 7 sowohl illustriert als auch erweitert und vertieft. Nach einer methodischen Einführung zu den vertiefenden Einzelfallanalysen (Kap. 7.1) und einem einleitenden Vergleich der formalen Gestaltung der Erwartungshorizonte (Kap. 7.2) ist das Kapitel – vor dem Hintergrund des komparativen Ansatzes der Studie – an den Bundesländern strukturiert, aus denen die diskutierten Fälle stammen (Kap. 7.3 Bayern, Kap. 7.4 Niedersachsen, Kap. 7.5 Thüringen). Die leitende Fragestellung der Fallanalysen fokussiert dabei auf den Maßstabsgehalt, den die in den Erwartungshorizonten gegebenen Qualitätsmerkmale aufweisen. Kapitel 7.6 gibt eine Zusammenfassung der fallorientierten Ergebnisse, welche in Kapitel 8 in eine Profilbildung der drei Bundesländer mündet. Diese Profilbildung arbeitet die Schwerpunktsetzungen der Bundesländer in Bezug auf Urteilsbildung in Aufgabenstellungen und Erwartungshorizonten vor dem Hintergrund der Analyseergebnisse interpretierend heraus.

Eine fragengeleitete Diskussion und Interpretation der Ergebnisse wird in Kapitel 9 entfaltet. Dabei sind drei Aspekte zentral: Zum einen wird gefragt, wie die in der Analyse zu Tage getretenen Differenzen der drei untersuchten Bundesländer zu bewerten sind (9.1). Es wird die These vertreten, dass zwischen Differenzen zu unterscheiden ist, die der religionssoziologischen Vielfalt Deutschlands Rechnung tragen und somit als kontextsensible Anpassungen zu

wahren sind, und jenen, die in Bezug auf die Vergleichbarkeit von Abschlüssen zu Ungerechtigkeit hinsichtlich der Vergabe von Bildungsqualifikationen führen und daher minimiert werden müssen. Zum anderen bildet die Einordnung der in den Erwartungshorizonten genannten Qualitätsmerkmale in den Forschungsdiskurs um Urteilsbildung einen weiteren zentralen Aspekt der Ergebnisdiskussion (9.2). Kapitel 9.3 reflektiert vor dem Hintergrund der Ergebnisse die Bedeutung der Bekenntnisorientierung des Religionsunterrichts im Urteilsbildungsprozess und fragt nach den Spezifika von Urteilsbildung im konfessionell gebundenen Religionsunterricht. In Kapitel 10 finden die Überlegungen mit einem Ausblick auf mögliche Anschlussforschungen einen resümierenden Abschluss.

2 Urteilskompetenz als komplexe Anforderung – Erkenntnisinteresse der Studie

Als Untersuchungsgegenstand der Studie werden in diesem einleitenden Kapitel ethische und religiöse Urteilsbildungen bestimmt, die normative Wertungen fordern (Werturteile). Sie werden als *kontrovers* hinsichtlich ihrer Fragestellungen und als *offen* hinsichtlich ihres Urteilsentscheides charakterisiert. Darüber hinaus wird eine Systematisierung der Fragestellungen religiöser und ethischer Urteilsbildung vorgenommen, welche die drei Kategorien »ethisch-philosophisch«, »ethisch-religiös« und »metaphysisch-religiös« umfasst. Dabei wird gezeigt, dass sich die Ausrichtung der Fragestellung auf die Urteilkategorien auswirkt, anhand derer geurteilt wird.

Nachdem der Urteilsgegenstand der Studie definiert wurde (Kapitel 2.1) dient Kapitel 2.2 der Klärung des der Studie zugrunde liegenden kompetenzorientierten Bildungsverständnisses. Dieses Bildungsverständnis formuliert Ansprüche an Lehre und Qualitätsüberprüfung, deren Herausforderungen an der Kompetenz Urteilsfähigkeit besonders deutlich werden. Kapitel 2.3 schließt an diese Herausforderungen an und skizziert die religionspädagogischen Positionen zum Umgang mit der Förderung und Bewertung von Einstellungen und Haltungen, die die normativen Aspekte der Urteilsbildung unmittelbar bedingen. Aus diesen beiden Kontexten – der Kompetenzorientierung und dem religionspädagogischen Diskurs über Einstellungen und Haltungen – leiten sich die Fragestellungen der Studie ab, welche in Kapitel 2.4 benannt werden.

2.1 Definition des Untersuchungsgegenstandes

Der Untersuchungsgegenstand der vorliegenden Studie ist die in den Einheitlichen Prüfungsanforderungen (EPA) für das Fach Evangelische Religionslehre geforderte Urteilsfähigkeit, die sich laut den Vorgaben auf »religiöse und ethische Fragen« bezieht, weshalb in der Studie auch mitunter von »ethischer und reli-

giöser Urteilskompetenz« gesprochen wird.¹ Mit den Adjektiven »religiös« und »ethisch« wird auf zwei Bereiche rekuriert, die sich zum einen klar voneinander unterscheiden lassen, zum anderen aber auch miteinander in einem Verhältnis stehen und Überschneidungen aufweisen können. Um die betrachteten Fragestellungen zu systematisieren, liegt der Arbeit eine kategoriale Unterscheidung zwischen »ethisch-philosophischen« Fragestellungen, »ethisch-religiösen« und »metaphysisch-religiösen« Fragestellungen zugrunde.

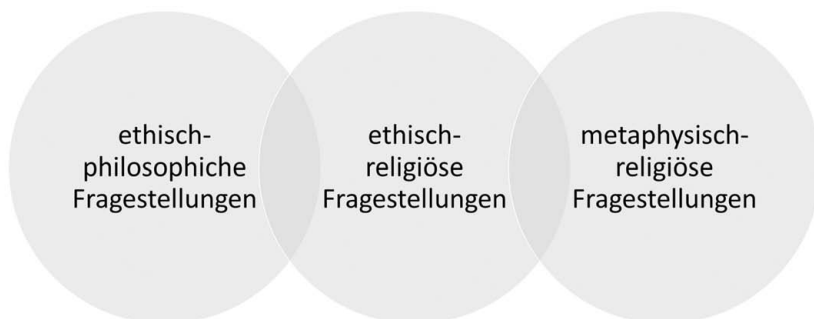


Abb. 1: Fragestellungen ethischer und religiöser Urteilsbildung.



Richten wir zunächst den Blick auf den Gegenstand ethischer Fragestellungen. Ethik lässt sich mit DIETMAR HÜBNER definieren als »die Wissenschaft von der Moral, d.h. diejenige Fachdisziplin, die sich damit befasst, welche Moralen es gibt, welche Begründungen sich für sie angeben lassen und welcher Logik ihre Begriffe, Aussagen und Argumentationen folgen.«²

Ethik hat folglich die *Moral* zum Gegenstand. Unter *Moral* wiederum »versteht man ein *Normensystem*, dessen Gegenstand *menschliches Handeln* ist und das einen Anspruch auf *unbedingte Gültigkeit* besitzt.«³ Das Normensystem wird von HÜBNER als eine »Sammlung von Maßstäben, Werten und Urteilen« genauer beschrieben.⁴ Er macht analog zu den drei Komponenten

¹ Vgl. EPA, 8.

Die Kurzzitierweise EPA bezieht sich hier und im Folgenden auf die EPA des Faches Evangelische Religionslehre. Wenn die EPA anderer Fächer zitiert werden, so ist dies kenntlich gemacht.

² DIETMAR HÜBNER, Einführung in die Philosophische Ethik. Göttingen 2014, 17.

³ A. a. O., 13. Hervorhebungen wie im Original.

⁴ Vgl. ebd.

seiner Definition drei wesentliche Merkmale des modernen Moralbegriffs deutlich: Erstens betont er, dass es eine Mehrzahl an Normensystemen gibt, die von verschiedenen Gruppen vertreten werden können, und Moral folglich problemlos im Plural stehen kann.⁵ Dabei drückt das Wort Moral an sich keine Wertung aus. Es bleibt also offen, ob der Sprecher der benannten Moral zustimmt oder nicht. Das Adjektiv »moralisch« kann neutral verwendet werden, wird aber häufig zur Wertung genutzt im Sinne von »aus Sicht der Moral der sprechenden Person richtig«⁶.

Zweitens stellt HÜBNER heraus, dass Moralen menschliches Verhalten in all seinen Komponenten regeln und drittens, dass die Moral nicht von kurzfristigen Zielsetzungen einer Person abhängig ist, sondern »unbedingte Gültigkeit« beansprucht.⁷ Der Unbedingtheitsanspruch der Moral macht diese also unabhängig von momentanen Wünschen und Vorhaben einer Person. Die Definition von Moral in der Enzyklopädie *Religion in Geschichte und Gegenwart* legt den Fokus auf die Interaktionsweisen, welche sich aus dem Normensystem ergeben und zu einer »relativen Stabilität« gelangen – also unbedingte Gültigkeit erheben.⁸ Dort heißt es:

»Der Ausdruck ›Moral‹ bezeichnet das Ganze einer in motivierenden und orientierenden Gewißeiten (Überzeugungen) und dadurch affektiver Strebens- bzw. Interessenlage samt zugehörigen Grundentschlüssen fundierten regelgemäßen, also zu relativer Stabilität gelangten Interaktionsweise, äquivalent mit ›Ethos‹ (gegenüber dem sozialen Akzent dieses Ausdrucks allenfalls mehr individuenbezogen); [...].«⁹

Ethik befindet sich im Verhältnis zur Moral auf einer Metaebene, was eine methodische Distanzierung zu dieser notwendig macht. Ethik zielt folglich auf eine wissenschaftliche Theorie über gutes Handeln, eingelebtes Ethos und vorhandener Moral ab. »[I]n ihr geht es um die theoretische Untersuchung beziehungsweise kritische Prüfung von gelebter Sittlichkeit und geltenden moralischen Normen.«¹⁰ Sie möchte eine wissenschaftsorientierte und analytische, also

⁵ Vgl. ebd.

⁶ A. a. O., 14.

⁷ Vgl. a. a. O., 15 f.

⁸ Vgl. EILERT HERMS, Art. Moral, in: *Religion in Geschichte und Gegenwart* ⁴2002 online. http://dx.doi.org.002dcbqo0d07.han.ulb.uni-jena.de/10.1163/2405-8262_rgg4_SIM_14450 (15.09.2020).

⁹ A. a. O.

¹⁰ HANS-RICHARD REUTER, Grundlagen und Methodik der Ethik, in: WOLFGANG HUBER/TORSTEN MEIREIS/HANS-RICHARD REUTER (Hrsg.), *Handbuch der Evangelischen Ethik*. München 2015, 9–123, hier 16. Im Original kursiv.

vordergründig auf kognitiver Ebene stattfindende und intersubjektiv nachvollziehbare Reflexion leisten.

Im Fokus der Ethik steht also das menschliche Handeln sowie die diesem zugrunde liegenden Werte, verschiedene Normensysteme und das Finden von Vorzugskriterien. Dabei besteht für die Reflexion der Moral, die in einer ethisch-moralischen Urteilsbildung stattfindet, kein Automatismus zu einem moralisch angemessenerem Verhalten.¹¹ Die Moral geht dem rationalen Nachdenken über sie voraus:

»Die Frage, warum wir überhaupt Interesse für unsere Mitmenschen aufbringen sollen, zeigt offenkundig bereits einen Mangel an etwas, das sich nicht rational einholen lässt, sondern Voraussetzung unserer rationalen Überlegungen sein muss, anders formuliert: was unsere rationalen Überlegungen immer schon bestimmen muss.«¹²

Evangelische Ethik fragt nach der »Prägekraft der christlichen Gewissheit«¹³; also danach, wie sich die aus dem Glauben resultierenden Wirklichkeitsannahmen – die christliche Gewissheit – auf die Lebensweise der Gläubigen auswirkt. Dabei ist zentral, dass die christlichen Perspektiven durchaus plural sind, worauf SVEND ANDERSEN eingeht, wenn er schreibt:

»Es wird die Meinung vertreten, der wesentlichste Unterschied [zwischen theologischer und philosophischer Ethik, Anm. K.M.] bestehe darin, dass die theologische Ethik von einer im Voraus gegebenen Auffassung vom rechten menschlichen Handeln ausgehe, nämlich derjenigen, die in der Bibel, in kirchlichen Bekenntnissen oder anderen Autoritäten enthalten ist. [...] Das ist jedoch bestenfalls eine Vereinfachung. [...] Demgegenüber kann behauptet werden, dass es für die Reflexion der theologischen Ethik ein bestimmtes Hauptthema gibt. Wir können es auf folgende Weise

¹¹ Vgl. MICHAEL ROTH, Ist Moral lehrbar? Überlegungen zur ethischen Kompetenz, in: BERND HARBECK-PINGEL/MICHAEL ROTH (Hrsg.), Emotionen und Verhalten in theologischer und philosophischer Perspektive. Leipzig 2012, 189–202, hier 190 f.

¹² A. a. O., 193.

¹³ EILERT HERMS, Art. Ethik V, in: Religion in Geschichte und Gegenwart ⁴1999 online. http://dx-1doi-1org-1002dcb8v0557.han.ulb.uni-jena.de/10.1163/2405-8262_rgg4_COM_04670 (19.09.2020). Die »christliche Gewissheit« expliziert HERMS als die »freimachende[] Gewißheit (Röm 8,38; Joh 8,32) über das Wesen des Schöpfers als Gnade und Wahrheit (Joh 1,14), somit auch über sein schöpferisches Wollen und Wirken als Liebe (Röm 8,38; 1 Joh 4,8), und die dem Geschenk des durch diese Gewißheit fundierten Freiseins durch den seiner Eigenart gemäßen freien Gebrauch entspricht (Gal 5,1 ff.) [sic]« (ebd.).

formulieren: Theologische Ethik ist die kritische Reflexion über diejenigen Auffassungen vom rechten menschlichen Handeln, die dem christlichen Glauben innewohnt.«¹⁴

Evangelische Ethik befasst sich ebenso wie die Philosophische Ethik mit dem gesamten Spektrum menschlichen Handelns. Sie hat jedoch eine spezifische Perspektive. Sie fragt danach, wie menschliches Handeln aus einer christlichen Sicht heraus zu bewerten ist.



ethisch-
philosophische
und
ethisch-religiöse
Fragestellungen

Abgeleitet von dem Selbstverständnis evangelischer Ethik können die Fragestellungen der Urteilsbildung im evangelischen Religionsunterricht *ethisch-philosophischer* Natur sein und keinen dezidierten Bezug zur Religion aufweisen, wie beispielsweise Fragestellungen nach der Aufnahme geflüchteter Menschen oder zu Möglichkeiten und Grenzen des technisch Machbaren.

Der Begriff *philosophisch* wird hier kontrastiv zu christlichen Konzeptionen der Weltdeutung verwendet. Diese Bedeutung ist bei PLATON angelegt, welcher den Begriff φιλοσοφία (*philosophía*) als einen spezifischen Terminus für eine neue Disziplinbeschreibung einführt, um das »dynamische[] Bemühen um Wissen im Bewusstsein eines Mangels an Wissen«, welches dem Menschen eigen ist, zu beschreiben.¹⁵ Dieses menschliche Bemühen steht in Abgrenzung zur σοφία (*sophía*), der Weisheit, die Gott zugeschrieben wird.¹⁶ Im heutigen Wortsinn meint Philosophie das »Streben nach Erkenntnis über den Sinn des Lebens, das Wesen der Welt und die Stellung des Menschen in der Welt«¹⁷.

Gleichzeitig rücken Fragestellungen in den Blick, welche von einer religiösen Weltsicht ausgehen oder religiöse Konzepte einschließen. Diese werden in der vorliegenden Studie als *ethisch-religiöse Fragestellungen* bezeichnet. Beispiele sind die Gegenwartsrelevanz eines lutherischen Freiheitsverständnisses oder die Bedeutung der Bergpredigt für die Politik. Charakteristisch für die theologische

¹⁴ Vgl. SVEND ANDERSEN, Einführung in die Ethik. Berlin/New York ²2005, 14 f. Hervorhebungen wie im Original.

¹⁵ Vgl. MICHAEL ERLER: Art. Philosophie I, in: Religion in Geschichte und Gegenwart ⁴2003 online. http://dx-1doi-1org-1002dcb8v0557.han.ulb.uni-jena.de/10.1163/2405-8262_rgg4_COM_024378 (15. 9.2020).

¹⁶ Vgl. ebd.

¹⁷ DUDENREDAKTION (Hrsg.), Art. Philosophie, in: Duden online. www.duden.de/rechtschreibung/Philosophie (15.09.2020).

Ethik und damit für beide Arten der Fragestellung ist die Perspektive, welche sich durch die dem Urteil zugrunde liegenden Urteilkategorien und Wertmaßstäbe zeigt.



Metaphysisch-
religiöse
Fragestellungen

Neben ethisch-religiösen Fragen existieren Fragestellungen, welche sich nicht in erster Linie auf menschliches Handeln, sondern auf göttliches Handeln beziehen, wie beispielsweise die Frage nach Gottesbildern, Sündenvergebung oder Auferstehungsvorstellungen. Urteilsbildungsprozesse in diesem Bereich werden im Folgenden als *metaphysisch-religiöse Fragestellungen* bezeichnet. Es wird gezeigt werden, dass diesen Fragestellungen

anders akzentuierte Urteilkriterien zugrunde liegen. So treten beispielsweise Urteilkategorien wie die Sozial- oder Umweltorientierung in den Hintergrund, während andere Kriterien wie die innere Kohärenz metaphysischer Vorstellungen in den Vordergrund rücken (vgl. Kapitel 4.2.3).

Die Studie betrachtet demnach Urteilsbildungen in ethisch-philosophischen, ethisch-religiösen und metaphysisch-religiösen Fragen, womit der Urteilsgegenstand definiert ist. Blicken wir nun auf die Urteilsbildung selbst:

Eine grundlegende Stufung von Urteilen stammt aus der Politikdidaktik. Eine dort gängige Unterscheidung führte PETER WEINBRENNER ein.¹⁸ Unterschieden wird zwischen konstatierenden Urteilen, welche sich auf die Beschreibung eines Sachverhaltes beziehen, deutenden Urteilen, die den Sachverhalt erklären und Werturteilen, in denen eine eigene Position in Bezug auf den Sachverhalt vertreten wird.¹⁹ Die von WEINBRENNER benannte Stufung eines Urteils hat sich als Unterscheidung zwischen *Sach- und Werturteil* breitenwirksam in vielen Fachdidaktiken durchgesetzt.²⁰ In den EPA des Faches Geschichte wird die Unter-

¹⁸ Vgl. PETER WEINBRENNER: Politische Urteilsbildung als Ziel und Inhalt des Politikunterrichts, in: BUNDESZENTRALE FÜR POLITISCHE BILDUNG (Hrsg.), Politische Urteilsbildung. Aufgaben und Wege für den Politikunterricht. Bonn 1997, 73–94, hier 74 f.

¹⁹ Vgl. a. a. O. sowie KARL-ERNST JEISMANN, der Urteilsbildung als einen Dreischritt beginnend mit einer Analyse über das Sachurteil zum Werturteil begreift (KARL-ERNST JEISMANN, Didaktik der Geschichte: Das spezifische Bedingungsfeld des Geschichtsunterrichts, in: GÜNTER C. BEHRMANN/KARL-ERNST JEISMANN/HANS SÜSSMUTH (Hrsg.), Geschichte und Politik. Didaktische Überlegungen eines kooperativen Unterrichts. Paderborn 1978, 50–107).

²⁰ In der Religionspädagogik hat sich diese Unterscheidung bisher nicht verankert, was sicher auch daran liegt, dass die religionspädagogische Debatte über Urteilsbildung von vornherein Werturteile und in diesem Zusammenhang Fragen nach der Wertebildung stark

scheidung anhand der Kriterien getroffen, die für einen Urteilsentscheid herangezogen werden. Dort heißt es:

»[...] Das Sachurteil beruht auf der Auswahl, Verknüpfung und Deutung historischer Sachverhalte innerhalb eines Bezugsrahmens, die zu einem Urteil führen. Gelungene Sachurteile weisen sich durch sachliche Angemessenheit, innere Stimmigkeit und ausreichende Triftigkeit von Argumenten aus. Darüber hinaus werden beim Werturteil ethische, moralische und normative Kategorien auf historische Sachverhalte angewendet und eigene Wertmaßstäbe reflektiert. Es werden Zeitbedingtheit bzw. Dauerhaftigkeit von Wertmaßstäben berücksichtigt.«²¹

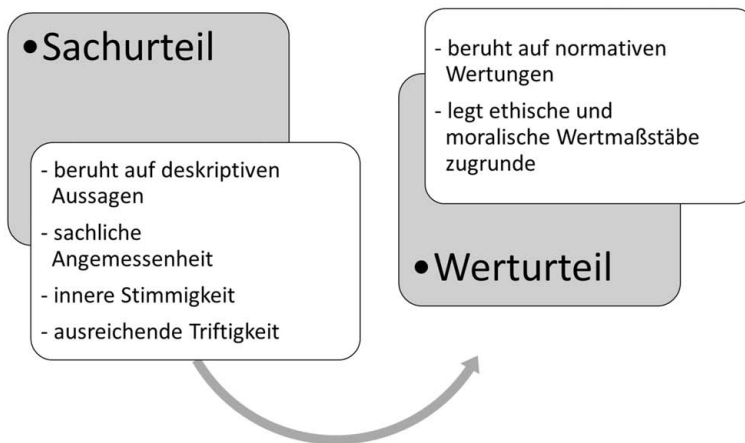


Abb. 2: Unterscheidung zwischen Sach- und Werturteil.

Ein fundiertes Sachurteil bildet die Grundlage für ein Werturteil. In der Begründung eines Werturteils spielen deskriptive und normative Aspekte zusammen: Während Urteile auf der sachlichen Ebene erläutert oder begründet werden müssen, findet die Rechtfertigung der Urteile auf der normativen Ebene des Werturteils statt.

Die Unterscheidung zwischen Sach- und Werturteil zielt auf die in der Ethik fest verankerte Unterscheidung zwischen *deskriptiven* und *normativen Aussagen*. Deskriptive also beschreibende Aussagen erheben einen Wahrheitsanspruch

fokussiert, während diese Unterteilung in anderen Fachdidaktiken, wie beispielsweise der Politik- oder Geschichtsdidaktik, zur Stärkung des Werturteils dient.

²¹ STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER (Hrsg.), Einheitliche Anforderungen für die Abiturprüfung im Fach Geschichte. Beschluss vom 1.12.1989 in der Fassung vom 10.2.2015, 4.

mit intersubjektiver Gültigkeit.²² Sie sind also überprüfbar, während normative Aussagen auf Normen und Werte rekurren, die keinen allgemeingültigen Wahrheitsanspruch haben, wenngleich intendiert ist, auch wertende Sätze intersubjektiv plausibel zu machen.²³ Beispielsweise ist die Aussage, dass durch gentechnisch veränderten Mais die Ernährungssicherung der Weltbevölkerung besser gewährleistet werden kann, da Ernteverluste durch Schädlinge, Unkraut oder Krankheitserreger auf diese Weise besser eingedämmt werden können, eine deskriptive Aussage. Von normativer Qualität ist die Aussage, dass es ein gutes und anzustrebendes Ziel ist, alle Menschen auf der Erde zu ernähren.²⁴ Man könnte folglich zu der Schlussfolgerung kommen, dass der Anbau von gentechnisch verändertem Mais zu befürworten sei. Fragen ethischer Urteilsbildungen zeichnen sich jedoch durch *Kontroversität* aus. Kontroversität beschreibt, dass sich auch gegensätzliche Argumente finden lassen, welche zu einem anderen Urteilsentscheid führen. So sind im Hinblick auf das Beispiel zum Anbau von gentechnisch verändertem Mais weitere deskriptive Aussagen zu berücksichtigen: Die Auswirkungen des Anbaus auf die Artenvielfalt und Ökosysteme sind bisher noch nicht hinreichend erforscht, wenngleich davon auszugehen ist, dass das Eingrenzen der züchterischen Vielfalt negative Auswirkungen haben wird.²⁵ Darüber hinaus ist zu erwarten, dass eine schnelle Resistenzentwicklung der zu bekämpfenden Schädlinge eintritt oder dass das Risiko neu auftretender Allergien beim Einsatz gentechnischer Zuchtmethoden steigt.²⁶ Allerdings gibt es auch hinsichtlich dieser Sachaussagen kontroverse Einschätzungen, wodurch die Fragestellung komplexer wird. Den Sachargumenten liegen konfligierende Werte zugrunde, die für ein Werturteil gegeneinander abgewogen werden müssen: So wären der Schutz der Ökosysteme und das Wohlbefinden unter allergischen Erkrankungen leidender Menschen gegen den Wert der eventuell nur kurzfristigen Ernährungssicherung abzuwägen.²⁷

²² Vgl. ANDERSEN, Einführung in die Ethik, 6.

²³ Vgl. ebd.

²⁴ Ausführlich zur ethischen Dimension von gentechnisch verändertem Mais siehe KATRIN PLATZER/KRISTINA SINEMUS, Ein Strukturmodell ethischer Urteilsbildung im Kontext moderner Biotechnologie und Gentechnik. Das Fallbeispiel Bt-Mais, in: CHRISTINE HAUSKELLER (Hrsg.), Wissenschaft verantworten. Soziale und ethische Orientierung in der technischen Zivilisation. Münster 2001, 91–107.

²⁵ Vgl. ebd.

²⁶ Vgl. ebd.

²⁷ Die Fragestellung nach dem Anbau von gentechnisch verändertem Mais würde nach der oben dargestellten Systematisierung als »ethisch-philosophische« Fragestellung kategorisiert, da sie keine unmittelbaren Bezüge zur Religion aufweist. Der Bezug zur Religion ergibt sich – wie oben erläutert – in diesen Fragen durch die Perspektivierung der Urteilstkatego-

Eine Positionierung – sofern sie nicht nur dem Schein nach eine Positionierung ist – setzt die Wahl zwischen mindestens zwei denkbaren, sich einander ausschließenden Optionen voraus.²⁸ Ohne die Existenz divergierender Werte würde eine Urteilsbildung obsolet. JÖRG KAISER und ULRICH HAGEMANN formulieren mit Blick auf den Geschichts- und Politikunterricht:

»Wird Urteilsbildung angestrebt, müssen prinzipiell mehrere begründete Stellungnahmen möglich sein: Urteile [...] sind zwar durch Thematisierung, didaktische Zugangsweisen und methodische Umsetzung in eine Richtung vorstrukturiert. Dennoch kann es nie um »die Lösung« in einer Urteilsphase gehen, sondern stets nur um »begründete Lösungsmöglichkeiten«.²⁹

Neben dem Charakteristikum der Kontroversität sprechen KAISER und HAGEMANN ein zweites konstitutives Charakteristikum an, nämlich die *Offenheit der Urteilsbildung*. Sie liegt darin begründet, dass die Kontroversität einer Fragestellung auf divergierenden Werten fußt, die jeweils begründbar sind. Werturteile bleiben daher stets widerlegbar und können nur im Rahmen der Akzeptanz der zugrunde liegenden normativen Setzungen Allgemeingültigkeit beanspruchen. Eine generelle Allgemeingültigkeit ist jedoch nicht möglich. Die Urteilsentscheidung muss dem Urteilenden obliegen – folglich »offen bleiben«.

Die Offenheit der Urteilsbildung ist hinsichtlich der oben eingeführten Differenzierung zwischen ethischen bzw. ethisch-religiösen Fragestellungen und metaphysisch-religiösen Fragestellungen noch einmal zu unterscheiden: Ethische und ethisch-religiöse Fragestellungen, die beispielsweise Schwangerschaftsabbrüchen oder die Relevanz des lutherischen Freiheitsverständnisses für die Gegenwart betreffen, können mittels deskriptiver Kategorien begründet und erläutert und anhand normativer Kategorien gerechtfertigt werden. Auch wenn das Werturteil nicht in den Kategorien »richtig« oder »falsch« bewertet werden kann, so kann es doch in seiner Begründung intersubjektive Plausibilität erfahren, wenngleich die Gültigkeit der Entscheidung die Akzeptanz der ihr zugrunde liegenden Prämissen voraussetzt.

Metaphysisch-religiöse Fragen, wie nach der Auferstehung der Toten oder nach Gottesbildern entziehen sich hingegen eines Sachurteils und normative

rien. Man würde folglich fragen, welche normativen Orientierungen sich aus einem christlichen Gottes- und Menschenverständnis für die Fragestellung ableiten lassen.

²⁸ Vgl. bspw. CHRISTIAN THEIN, Problemreflexion und Urteilsbildung im Philosophieunterricht, in: Mitteilungen des Fachverbandes Philosophie e.V., 56 (2016), 85–93, hier 88. Abrufbar unter http://fv-philosophie-nrw.de/wp-content/uploads/2016/12/Mitteilungen_gesamt_2016_web.pdf (03.07.2019).

²⁹ Vgl. JÖRG KAISER/ULRICH HAGEMANN, Urteilsbildung im Geschichts- und Politikunterricht. Berlin 2010, 38.

Kategorien wie Werte führen nur bedingt weiter. Sie sind eng an die Wirklichkeitsannahmen und Glaubensüberzeugungen einer Person geknüpft: Moralische Pflichten und Normen wirken sich nicht – wie es in nicht-metaphysischen Fragestellungen der Fall ist – auf das Konstrukt der eigenen Weltdeutung und die persönliche Moral aus, wie folgendes Beispiel illustrieren kann: Ich glaube nicht an eine göttliche Auferweckung nach dem Tod, weil ich mich an moralische Pflichten gebunden weiß. Das Gegenteil ist der Fall. Die eigene Weltdeutung wirkt sich gegebenenfalls auf moralische Pflichten und Normen aus: Weil ich an die göttliche Auferweckung nach dem Tod glaube, verpflichte ich mich im Leben moralischer Grundsätze. Dies zieht Unterschiede in Hinblick auf die Allgemeingültigkeit der Urteilsentscheide nach sich: Während ein ethisches, nicht-metaphysisches Urteil den Anspruch hat, einen Entscheid begründen zu können, der für eine größtmögliche Gruppe gelten sollte und der andere Lösungen begründet ablehnt, bleiben Urteilsentscheide im Bereich metaphysischer Fragen auf den Personenkreis begrenzt, der das Konstrukt teilt. Vor diesem Hintergrund kann THERESA SCHWARZKOPF für metaphysisch-religiöse Fragen »Mehrperspektivität« als »das gleichberechtigte Nebeneinander von Antwortmöglichkeiten«³⁰ definieren, während »Mehrperspektivität« in ethischen, nicht-metaphysischen Fragestellungen die Wahrnehmung anderer Positionen, aber nicht deren Gleichberechtigung bedeuten kann.³¹ Die innere Kohärenz der eigenen Konstrukte und die gleichberechtigte Akzeptanz anderer Antwortmöglichkeiten erhalten somit in Bezug auf metaphysisch-religiöse Fragestellungen eine größere Bedeutung als es in nicht-metaphysischen Fragen der Fall ist.

Zusammenfassend lässt sich der Gegenstand der Studie wie folgt bestimmen: Untersucht werden Urteilsbildungen, welche sich durch eine normative Dimension auszeichnen (Werturteile) und durch die Charakteristika *Kontroversität der Fragestellung* und *Offenheit des Urteilsentscheides* bestimmt sind. Obwohl sich in Ausnahmefällen auch religiöse oder ethische Fragestellungen finden lassen, welche durch ein reines Sachurteil bewertet werden können, werden in der Studie nur Fragestellungen normativer Urteilsbildung untersucht, also jene, die auch ein Werturteil fordern, »etwa in Gestalt einer Verteidigung oder eines Abweisens von bestimmten Normen«.³² HEINZ VON FOERSTER definiert diese Fragestellungen als *unentscheidbare Fragen*, da sie »durch die Freiheit unserer Wahl bestimmt« sind und wir für sie prinzipiell die Verantwortung tragen

³⁰ THERESA SCHWARZKOPF, *Vielfältigkeit Denken. Wie Schülerinnen und Schüler im Religionsunterricht argumentieren lernen*. Stuttgart 2016, 181.

³¹ Die Studie THERESA SCHWARZKOPFS zur Argumentationskompetenz in diesen metaphysischen Fragen, welche sie als »unentscheidbare Fragen« bezeichnet, wird in Kapitel 4.2.3 vorgestellt. Eine Abgrenzung zu ethischen Fragen nimmt sie nicht vor.

³² Vgl. ANDERSEN, *Einführung in die Ethik*, 9.

müssen.³³ Zu begründen ist die Fokussierung auf diese unentscheidbaren Fragen durch die Herausforderungen dieses Lernbereichs, der sich

»weniger durch ›richtig‹ oder ›falsch‹ verifizierbare Ergebnisse des Lernens abbilden und erfassen lässt, sondern sich durch eine problematisierende sowie reflektierende Auseinandersetzung mit moralischen Konfliktfällen auszeichnet und damit auf reflexive Denkmodi setzt, die nicht notwendigerweise zu abschließend als eindeutig richtig auszuweisenden Ergebnissen führen.«³⁴

Das Bewerten der angesprochenen reflexiven Denkmodi stellt besondere Herausforderungen an die Objektivität, die durch den Einfluss subjektiver Einstellungen, Haltungen und Werte auf den Urteilsentscheid zustande kommt. Diese Herausforderungen sollen im Folgenden im Kontext eines kompetenzorientierten Bildungsverständnisses herausgearbeitet werden.

2.2 Verortung der Studie in einem kompetenzorientierten Bildungsverständnis

2.2.1 Zentrale Aspekte eines kompetenzorientierten Bildungsverständnisses

Der Studie liegt ein kompetenzorientiertes Bildungsverständnis zugrunde. Dies ist angesichts der Ausrichtung aller bildungsadministrativen Vorgaben an Kom-

³³ Vgl. HEINZ VON FOERSTER, Lethologie. Eine Theorie des Erlernens und Erwissens angesichts von Unwißbarem, Unbestimmbarem und Unentscheidbarem, in: REINHARD VOB (Hrsg.), Die Schule neu erfinden. Systematisch-konstruktivistische Annäherungen an Schule und Pädagogik. Neuwied ³1999, 14–32, hier 29.

Die Psychologie und Sozialforschung bezeichnet Urteilsbildungsprozesse zu diesen Fragen als *moralische Urteilsbildung* (vgl. Kap. 4.1.2).

Inwieweit menschliches Handeln tatsächlich einer freien Wahl von Optionen entspringt, wurde in Philosophie und Theologie lange diskutiert und spielt bis heute für jedes Konzept der Weltbetrachtung eine zentrale Rolle. Auf diese Diskurse um den Freiheitsbegriff sei hier jedoch nur verwiesen. Urteilsbildung setzt die Annahme einer Willens- und Wahlfreiheit in gewissem Maße voraus.

³⁴ Vgl. MARTINA VON HEYNITZ, Die Entwicklung von Testaufgaben zur Erhebung ethisch-moralischer Grundkenntnisse, Urteils- und Handlungsentwurfskompetenz, in: DIETRICH BENNER/ROUMIANA NIKOLOVA (Hrsg.), Ethisch-Moralische Kompetenz als Teil öffentlicher Bildung. Der Berliner Ansatz zur Konstruktion und Erhebung ethisch-moralischer Kompetenzniveaus im öffentlichen Erziehungs- und Bildungssystem mit einem Ausblick auf Projekte zu ETiK-International. Paderborn 2016, 69–100, hier 69.

petenzen zum einen naheliegend, zum anderen jedoch auch angesichts der vielschichtigen Skepsis gegenüber der Kompetenzorientierung – gerade im religionspädagogischen Diskurs – begründungsbedürftig:

Unter dem Begriff der Kompetenz versammeln sich viele unterschiedliche Konzepte aus unterschiedlichen Professionen, von denen wiederholt betont wird, dass sie kaum eine Synthese hinsichtlich ihres Kompetenzverständnisses zulassen.³⁵ Als konstitutiv für die im Bildungsbereich diskutierten Konzeptionen kann über die unterschiedlichen Ausprägungen der Konzeptionen hinweg jedoch die Handlungsrelevanz der zu erlernenden Fähigkeiten benannt werden.³⁶ Als eine grundlegende Differenzierung lassen sich im pädagogischen Kontext nach NICLAS SCHAPER konzeptionell zwei psychologisch fundierte Kompetenzverständnisse unterscheiden³⁷:

1. Zum einen existiert ein pädagogisch-psychologischer Kompetenzbegriff, welcher auf der Kompetenzdefinition FRANZ E. WEINERTS beruht, kognitionspsychologische Theorien von Leistungsvoraussetzungen zugrunde legt und häufig zwischen kognitiven Kompetenzdimensionen und nicht-kognitive Kompetenzdimensionen unterscheidet, wobei die kognitiven Dimensionen in den Fokus der Debatte rücken. Kompetenzen werden dabei domänenspezifisch verstanden als »die bei Individuen verfügbaren oder erlernbaren kognitiven Fähigkeiten und Fertigkeiten, bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösungen in variablen

³⁵ Vgl. RÜDIGER PREIßER/REINHARD VÖLZKE, Kompetenzbilanzierung. Hintergründe, Verfahren, Entwicklungsnotwendigkeit, in: REPORT 30 (2007), 62–71, hier 62. Abrufbar unter: <https://www.die-bonn.de/doks/preisser0701.pdf> (14.6.2020); MARISA KAUFHOLD: Kompetenz und Kompetenzerfassung. Analyse und Beurteilung von Verfahren der Kompetenzerfassung. Wiesbaden 2006, 21 f.; DOMINIK HELBLING, Religiöse Herausforderung und religiöse Kompetenz. Empirische Sondierungen zu einer subjektorientierten und kompetenzbasierten Religionsdidaktik. Wien/Zürich 2010, 54–86.

³⁶ Vgl. PREIßER/VÖLZKE, Kompetenzbilanzierung, 62.

DOMINIK HELBLING, der sechs verschiedene pädagogische Kompetenzansätze miteinander vergleicht, benennt die Autonomie, Handlungs- und Wertorientierung als drei übereinstimmende Merkmale der von ihm untersuchten Konzeptionen (vgl. HELBLING, Religiöse Herausforderung, 84).

³⁷ Vgl. NICLAS SCHAPER, Ansätze zur Kompetenzmodellierung und -messung im Rahmen einer kompetenzorientierten Hochschuldidaktik, in: FLORIAN BRUCKMANN/OLIVER REIS/MONIKA SCHEIDLER (Hrsg.), Kompetenzorientierte Lehre in der Theologie. Konkretion – Reflexion – Perspektiven. Berlin 2011, 37–62, hier 39 f.

Situationen erfolgreich und verantwortungsvoll nutzen zu können«.³⁸ Da sich die sogenannte Klieme-Expertise auf dieses Kompetenzverständnis stützt, ist die Kompetenzdebatte im schulischen Kontext wesentlich an der WEINERT'schen Definition ausgerichtet. Kompetenzen werden in dieser Tradition als Persönlichkeitsdispositionen verstanden, welche »als kognitive Wissenskonstrukte wirksam werden (= Handlungspotentiale)«.³⁹

2. Parallel dazu hat sich ein Kompetenzverständnis entwickelt, welches in erster Linie aus dem berufspädagogischen Kontext stammt und sich an Theorien der Handlungsregulation orientiert. Kompetenzen werden in diesem Verständnis als »erlernbare Fähigkeiten für einen objektiv beschreibbaren Tätigkeitsprozess«⁴⁰ im Sinne von Arbeitstechniken verstanden. Berufsrelevante und andere Kompetenzen werden dabei durch die Bereiche Fach-, Methoden-, Sozial-, und Personal- bzw. Selbstkompetenz in allgemeiner Form strukturiert und anhand berufsfeldspezifischer Komponenten ausdifferenziert.⁴¹ Dadurch ergibt sich, so SCHAPER, eine hohe Vergleichbarkeit und Anschlussfähigkeit zwischen den Kompetenzmodellen, in der Praxis jedoch auch häufig eine pragmatische und wenig strukturierte Ausformulierung der einzelnen Kompetenzen.⁴² Problematisch ist an dieser Orientierung an berufsnahen, objektivistischen Erwartungsrahmen, dass bei diesen vorausgesetzt wird, dass die Dimensionen eigenständig und getrennt voneinander als Konstrukte entwickelt und leicht auf andere Situationen transformiert werden können.⁴³ Gerade im Kontext vieler Lehrpläne ist auch dieses Kompetenzverständnis konzeptionsbildend in das System Schule aufgenommen worden.

Der vorliegenden Studie liegt ein pädagogisch-psychologisches Kompetenzverständnis zugrunde, welches Dimensionen von Kompetenzen nicht unabhängig voneinander fasst, sondern darauf abzielt, in komplexen Problemsituationen handlungsfähig zu sein. Für dieses Bildungsverständnis ist zentral, dass unter-

³⁸ FRANZ E. WEINERT, Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit, in: DERS. (Hrsg.), Leistungsmessungen in Schulen. Weinheim 2001, 17–31, hier 27.

Zur Kontextualisierung der Definition WEINERTS in der Klieme-Expertise vgl. BMBF (Hrsg.), Bildungsstandards, 72.

³⁹ OLIVER REIS, Sinn und Umsetzung der Kompetenzorientierung – Lehre »von hinten« denken, in: PATRICK BECKER (Hrsg.), Studienreform in der Theologie. Eine Bestandsaufnahme. Berlin 2011, 108–127, hier 108.

⁴⁰ A. a. O., 108.

⁴¹ Vgl. SCHAPER, Kompetenzmodellierung, 39 f.

⁴² Vgl. a. a. O., 40.

⁴³ Vgl. REIS, Kompetenzorientierung, 111.

32 2 Urteilskompetenz als komplexe Anforderung

schiedliche Facetten der Persönlichkeit angesprochen werden, die die Bewältigung von Anforderungssituation bedingen.⁴⁴ Aus dem hier vertretenen kompetenzorientierten Bildungsverständnis heraus entsteht Bildung dann, wenn Individuen Inhalte durch eine Tätigkeit bearbeiten und in diesem Prozess eine Haltung zum Inhalt ausbilden.⁴⁵

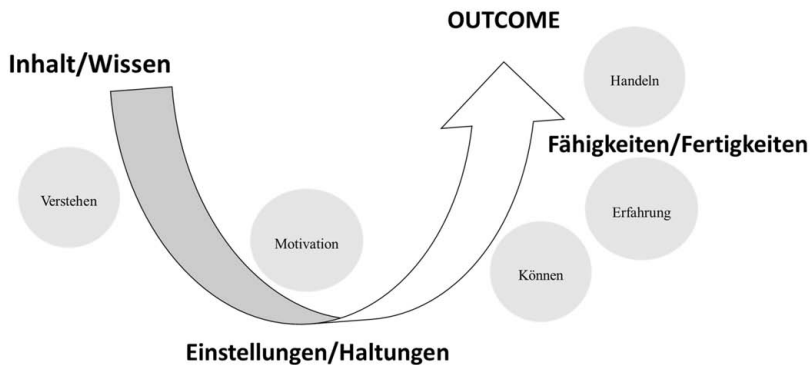


Abb. 3: Schematische Darstellung eines kompetenzorientierten Bildungsverständnisses.

Kompetenzen sind dabei nur dann nötig, »wenn Fertigkeiten nicht mehr ausreichen, weil die Komplexität der Situation keine Lösungstechnik mehr zulässt.«⁴⁶ Daraus folgt, dass das Subjekt seinen eigenen Lösungsweg finden muss, und dadurch der Situation transformativ begegnet.⁴⁷

Vor dem Hintergrund der zentralen Kritik an der Kompetenzorientierung, welche sich auf

- Ökonomisierung und Funktionalisierung von Bildung,
- Technokratisierung von Lernprozessen,
- Reduktion von Komplexität,
- einseitige Betonung von Testleistungen und

⁴⁴ Vgl. BMBF (Hrsg.), Bildungsstandards, 73; HELBIG, Religiöse Herausforderung, 85.

⁴⁵ REIS und SCHWARZKOPF schreiben: »[A]us Sicht der Kompetenzorientierung [ist] der Gegenstand immer ein Inhalt, der untrennbar mit einer bestimmten Tätigkeit bearbeitet wird«, und führen in Bezug auf die Religionspädagogik weiter aus: »Es wird nun für die religiösen Lernprozesse entscheidend darauf ankommen, die Inhalte mit den Tätigkeiten so zu koppeln, dass sinnvolle Lernziele entstehen, die genau die Freiheit im religiösen Lernen wahren und trotzdem normativ erwartbare Handlungsformen darstellen.« (OLIVER REIS/THERESA SCHWARZKOPF, Diagnose religiöser Lernprozesse, 46.)

⁴⁶ REIS, Kompetenzorientierung, 113.

⁴⁷ Vgl. a.a.O., 112 f.

- das »Abarbeiten von Stufenmodellen und Checklisten« zu Lasten eines tiefen Inhaltsverständnisses

bezieht, ist die konzeptionell in der Kompetenzorientierung verankerte Subjektorientierung des pädagogisch-psychologischen Kompetenzverständnisses nochmals hervorzuheben, welche sich auf die ganze Person ausrichtet – auf Wissen, Fähigkeiten, Verstehen, Können, Handeln, Erfahrung, Motivation.⁴⁸ Ziel einer so verstandenen Kompetenzorientierung ist, dass

»Subjekt[e] befähigt werden, den unberechenbaren Anforderungen des Lebens begegnen zu können, indem [sie] [...] die notwendigen Komponenten zur Bewältigung einer Situation abrufen, kombinieren und angemessen umsetzen [können]. Handlungsfähigkeit meint folglich das Vermögen, unbekannte und unerwartete Probleme lösen zu können.«⁴⁹

Eine Zergliederung, Formalisierung und Technokratisierung von Lernprozessen hingegen lässt sich ebenso wie die Reduktion von Komplexität als Fehlentwicklungen der Kompetenzorientierung bestimmen, die in einem pädagogisch-psychologischen Kompetenzverständnis nicht intendiert ist. Hinsichtlich der Ganzheitlichkeit des Lernens ist die konzeptionelle Grundausrichtung von Kompetenzen an ein klassisches Bildungsverständnis anknüpfungsfähig, welches mit WILHELM VON HUMBOLDT Bildung als »die Verknüpfung unseres Ichs mit der Welt« und mit WOLFGANG KLAFFKI Bildung »als Befähigung zu vernünftiger Selbstbestimmung, die die Emanzipation von Fremdbestimmung voraussetzt oder einschließt, als Befähigung zur Autonomie, zur Freiheit eigenen Denkens und eigener moralischer Entscheidungen« definiert.⁵⁰ KLAFFKI hebt noch einmal

⁴⁸ Zur allgemeinpädagogischen Kritik an der Kompetenzorientierung vgl. u. a. ECKHARD KLIEME/HANS BRÜGELMANN, »Standards vorgeben?« Pro und Contra, in: Pädagogik 56 (2004), 50–51; KASPAR SPINNER, Der standardisierte Schüler, in: Informationen zur Deutschdidaktik 31 (2007), 130–133; JENS BELJAN, Schule als Resonanzraum und Entfremdungszone. Eine neue Perspektive auf Bildung. Weinheim/Basel ²2019, 7–29.

Da das Kapitel lediglich eine Verortung der Studie zum Ziel hat, soll die Kritik an der Kompetenzorientierung zwar reflektiert, jedoch nicht weiter erläutert werden. Eine Aufarbeitung der Kompetenzdebatte wurde in der bereits zitierten Dissertationsschrift von DOMINIK HELBLING geleistet (vgl. DERS., Religiöse Herausforderung, insbesondere 43–50). Eine Auseinandersetzung mit der spezifisch religionspädagogischen Kritik an der Kompetenzorientierung erfolgt im folgenden Abschnitt (vgl. Kap. 2.3).

⁴⁹ HELBLING, Religiöse Herausforderung, 85.

⁵⁰ WILHELM VON HUMBOLDT, Theorie der Bildung des Menschen, in: ANDREAS FLINTNER/KLAUS GIEL (Hrsg.), Wilhelm von Humboldt. Werke in fünf Bänden. Erster Band: Schriften zur Anthropologie und Geschichte. Berlin 1960, 236.

pointiert hervor, was auch die Kompetenzorientierung unterstreicht, nämlich dass die »Selbsttätigkeit die zentrale Vollzugsform des Bildungsprozesses« sei.⁵¹ Materiale und formale Aspekte sind nach KLAFFKI in einem Bildungsbegriff zu vereinen – eine Forderung, die auch die Kompetenzorientierung erhebt.⁵²

Was sich entschieden von einem klassischen Bildungsverständnis unterscheidet, ist, dass sich die Kompetenzorientierung dezidiert nicht als Selbstzweck versteht. Es geht ihr folglich um die Abkehr von »totem Wissen« hin zu einer handlungsorientierten Einbindung von Wissen in komplexe lebensweltbezogene Anforderungssituationen. Intendiert wird dabei, den Wissensaufbau bezüglich einer Kompetenz konkret in Niveaustufen beschreiben und damit gezielter anleiten und messen zu können. Indem die Kompetenzorientierung dafür plädiert, den sogenannten *Outcome*, also den Ertrag des Lernprozesses, klar zu definieren, geht sie davon aus, dass für die Kernbestände des in der Schule zu Erlernenden operationalisier- und messbare Anforderungen benannt werden können. Zugleich beansprucht sie, den schrittweisen Aufbau dieses Outcomes beschreiben und dadurch für alle nachvollziehbar und prinzipiell erlernbar machen zu können.⁵³

Zusammenfassend kann festgehalten werden, dass das von bildungsadministrativer Seite vorgegebene Paradigma der Kompetenzorientierung – in erster Linie aus pädagogischer Sicht – nicht unumstritten ist. Es wurde die These dargelegt, dass in Bezug auf die Kritikpunkte an der Kompetenzorientierung unterschieden werden muss zwischen tatsächlich im Konzept der Kompetenzorientierung verankerten Schwachstellen und vom Konzept selbst nicht intendierten »Fehlwüchsen«. Diese intensive Auseinandersetzung mit den Ansprüchen der Kompetenzorientierung ist von wesentlicher Bedeutung, um die Ergebnisse der Analyse von Abituraufgaben einordnen und bewerten zu können. Leistungsbeurteilungen und Bewertungskriterien für Urteilskompetenz müssen sich von den Kritikpunkten wie der Zergliederung und Formalisierung von Bildung, aber auch die Reduktion von Komplexität hinterfragen lassen. Die An-

⁵¹ Vgl. WOLFGANG KLAFFKI, Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch konstruktive Didaktik. Weinheim/Basel ⁶2007, 19.

⁵² KLAFFKI bringt dieses wechselseitige Verhältnis beider Aspekte in der vielzitierten Formel zusammen: »Bildung ist Erschlossensein einer dinglichen und geistigen Wirklichkeit für einen Menschen – das ist der objektive oder materiale Aspekt; aber das heißt zugleich: Erschlossensein dieses Menschen für diese seine Wirklichkeit – das ist der subjektive der formale Aspekt« (DERS., Studien zur Bildungstheorie und Didaktik. Weinheim/Basel 1975, 45).

⁵³ In diesem Punkt unterscheidet sich die Kompetenzorientierung von einem resonanzpädagogischen Bildungsverständnis, welches die Unverfügbarkeit von Lernprozessen akzentuiert (vgl. BELIAN, Schule als Resonanzraum).

sprüche der Kompetenzorientierung werden im Folgenden aus diagnostischer Perspektive nochmals spezifiziert.

2.2.2 Der Konstruktcharakter von Kompetenzen

Aus diagnostischer Perspektive ist das Wissen um den Konstruktcharakter der zu prüfenden Kompetenz von zentraler Bedeutung: Kompetenzen lassen sich nicht unmittelbar messen, da sie nicht gegenständlich und dadurch nicht unmittelbar zugänglich sind wie beispielsweise die Körpergröße oder das Gewicht eines Menschen. Es handelt sich daher um Konstrukte. »Konstrukte werden aus theoretischen Zusammenhängen heraus erschlossen und sind somit Bestandteile von theoretischen Aussagen. Die Ausprägung eines Konstrukts kann nur indirekt aus messbaren Indikatoren geschlossen werden.«⁵⁴ Um ein Konstrukt messen zu können, bedarf es einer möglichst präzisen Bestimmung des Konstrukts. Diese muss durch Operationalisierung mit Indikatoren verknüpft sein, welche Informationen zur Ausprägung des Konstrukts geben.⁵⁵ Für eine Messung ist darüber hinaus die Zuordnung von Objekten zu Zahlen zentral.⁵⁶ Die Definition und Präzisierung der Konstrukte ist dynamisch und vom wissenschaftlichen Erkenntnisfortschritt abhängig.⁵⁷ Testtheoretisch gilt in Bezug auf die Steigerung der Messgüte, dass eine präzise Konstruktvorstellung sowie ein hoher Präspezifikationsgrad, der durch die vorherige Zuordnung möglicher Lösungen zu verschiedenen Anforderungsniveaus entsteht, die Validität einer Prüfung steigern (vgl. dazu Kapitel 3.3.2).⁵⁸ Mit Blick auf Urteilsfähigkeit stellt sich die Frage, wie das Konstrukt der Kompetenz auf Grundlage des wissenschaftlichen For-

⁵⁴ DETLEF URHAHNE/MARKUS DRESSSEL/FRANK FISCHER (Hrsg.), *Psychologie für den Lehrberuf*. Berlin/Heidelberg 2019, 622. Vgl. dazu auch a. a. O., 537 f.

⁵⁵ MICHAEL EID et al. definieren Operationalisierung als Messbarmachung von Konstrukten durch empirisch fassbare und quantifizierbare Größen (vgl. MICHAEL EID/MARIO GOLLWITZER/MANFRED SCHMITT, *Statistik und Forschungsmethoden*. Weinheim/Basel ⁵2017, 42).

⁵⁶ In Bezug auf das »Messen« ist zentral, dass »die Zuordnung von Zahlen zu Objekten nach bestimmten Regeln [gewährleistet] [...], dass bestimmte (interessierende) Relationen in der Menge der Objekte in der Menge der Zahlen erhalten bleiben.« (vgl. EID/GOLLWITZER/SCHMITT, *Statistik*, 106).

⁵⁷ Vgl. JÖRG JOST/INGRID BÖTTCHER, *Leistungen messen, bewerten und beurteilen*, in: MICHAEL BECKER-MROTZEK/INGRID BÖTTCHER (Hrsg.), *Schreibkompetenz entwickeln und beurteilen*. Berlin ⁵2014, 113–144, hier 115.

⁵⁸ Vgl. MARC WORBACH/BARBARA DRECHSEL/KLAUS H. CARSTENSEN, *Messen und Bewerten von Lernergebnissen*, in: DETLEF URHAHNE/MARKUS DRESSSEL/FRANK FISCHER (Hrsg.), *Psychologie für den Lehrberuf*. Berlin/Heidelberg 2019, 493–516, hier 500.

schungsstandes zu beschreiben ist und welches Konstrukt den aktuellen Prüfungsvorgaben und -konventionen zugrunde liegt.

OLIVER REIS und THERESA SCHWARZKOPF folgend können Kompetenzen als multimodale Konstrukte beschrieben werden, die die Bestandteile »Wissen, Fertigkeiten, Fähigkeiten, Motive und Einstellungen/Haltungen« beinhalten.⁵⁹ Lernprozesse, welche unter diesem Komplexitätsniveau verlaufen, weil sie nicht das gesamte multimodale Konstrukt erfordern, werden von REIS und SCHWARZKOPF nicht als kompetenzorientiert im engeren Sinne betrachtet.⁶⁰ Dieses Verständnis von Bildung, das die Einstellungen und Haltungen untrennbar mit den Fähigkeiten und Fertigkeiten Wissen für »Problemlösungen in variablen Situationen« zu nutzen, verknüpft, hat zur Konsequenz, dass auch die Ausbildung von Einstellungen und Haltungen zum Lerngegenstand zu einem wesentlichen Bestandteil von Bildung und damit zum erwünschten Outcome wird. Dies ist mit Blick auf Werturteile, welche eine normative Bewertung der Fragestellung fordern von besonderer Relevanz, da das Abwägen divergierender Normen letztlich auf den subjektiven Einstellungen und Haltungen des Urteilenden basiert (vgl. Kapitel 2.1).⁶¹ Da die Kompetenzorientierung den Anspruch hat, Niveaustufen der Kompetenzentwicklungen und die Outcomes möglichst präzise zu beschreiben und vor allem messbar und operationalisierbar zu machen, stellt sich die Frage, inwieweit dies auch in Bezug auf Einstellungen und Haltungen zu den Lerngegenständen möglich ist.⁶² Die Asymmetrien zwischen Lehrenden und

⁵⁹ Vgl. REIS/SCHWARZKOPF, Diagnose religiöser Lernprozesse, 57 f.

⁶⁰ Vgl. a. a. O., 58 (Formatierungen wurden nicht übernommen). Als kompetenzorientiert im weiteren Sinne definieren SCHWARZKOPF und REIS einen Lernprozess, »wenn er mittelfristig einen kompetenzbezogenen Lernprozess vorbereitet« (vgl. a. a. O.; Formatierungen wurden nicht übernommen).

⁶¹ Der US-amerikanische Psychologe JONATHAN HAIDT vertritt die These, dass Urteilen in erster Linie auf affektive und emotionale Aspekte zurückzuführen ist und dass rationale Erklärungsmuster sekundär sind. Er stellt damit in Frage, dass einem Werturteil ein vorheriges Abwägen unterschiedlicher Fakten, also ein Sachurteil, vorausgeht (vgl. JONATHAN HAIDT, The Emotional Dog and its Rational Tail. A Social Intuitionist Approach to Moral Judgment, in: Psychological review 108 [2001], 814–834. Abrufbar unter: https://www.motherrjones.com/wp-content/uploads/emotional_dog_and_rational_tail.pdf [09.08.2018]). Der Befund HAIDTS kann didaktisch aufgegriffen werden, indem man den ersten »Bauchentscheidungen« der Schülerinnen und Schüler Raum gibt und von diesen ausgehend den Urteilsbildungsprozess entwickelt. Vgl. bspw. die intuitionistische Lehrstückdidaktik, wie sie MARIO ZIEGLER vertritt (MARIO ZIEGLER, Die festgefügteten Urteile des moralisch Naiven. Möglichkeiten der Reflexion und Revision am Beispiel von Kurosawas Film »Rashomon«, in: MARKUS TIEDEMANN [Hrsg.], Werte und Wertevermittlung. Dresden 2018, 63–73).

⁶² In der Klieme-Expertise wird dies am Beispiel der Fremdsprachenkompetenz offensichtlich, welche laut der Expertise die Anforderung einschließt, sich »in der Intention und

Lernenden, die beispielsweise durch entwicklungspsychologische Unterschiede, aber auch durch Abhängigkeits- und Machtverhältnisse vorhanden sind, stellen ein wesentliches Argument gegen die Ausbildung spezifischer Einstellungen und Haltungen dar, auch dann, wenn man für diese gute Gründe zu haben meint. Diese Einsicht wurde im bis heute breit rezipierten Beutelsbacher Konsens der Politikdidaktik 1976 in einem »Überwältigungsverbot«, einem »Kontroversitätsgebot« und in der Forderung nach einer Befähigung zur Mündigkeit der Schülerinnen und Schüler festgehalten.⁶³ Die Ethik- und Philosophiedidaktik verpflichtete sich im Dresdner Konsens 2016 ebenso, nicht »durch die Betonung einer Sichtweise« in die Urteilsbildung der Lernenden einzugreifen.⁶⁴

Da Religiosität sehr eng mit individuellen Einstellungen und Haltungen verbunden ist und da der Religionsunterricht durch seine konfessionelle Bindung für spezifische Einstellungen und Werte steht, führte die Implementierung eines kompetenzorientierten Wissensverständnisses auch in der Religionspädagogik zu einer breiten Debatte über eben die Frage, ob die anzustrebenden Outcomes des Religionsunterrichts spezifische Einstellungen und Haltungen umfassen sollten und damit verbunden zu der prinzipiellen Frage nach der Lehr- und Bewertbarkeit von Einstellungen und Haltungen. Diese Debatte wird im folgenden Kapitel aufgearbeitet, um aus dem Forschungsdiskurs die zentralen Fragestellungen der Studie abzuleiten.

Motivation [...] offen und akzeptierend mit anderen Kulturen auseinander zu setzen« (vgl. BMBF [Hrsg.], Bildungsstandards, 73).

Die Ausbildung einer Einstellung zum Lerngegenstand wird in diesem Zusammenhang deutlich als Ziel eines kompetenzorientierten Lernens definiert, wobei in diesem Beispiel sogar eine spezifische Einstellung zum Gegenstand gefordert ist – nämlich eine offene und akzeptierende.

⁶³ HANS-GEORG WEHLING, Konsens á la Beutelsbach? Nachlese zu einem Expertengespräch, in: SIEGFRIED SCHIELE/HERBERT SCHNEIDER (Hrsg.), Das Konsensproblem in der politischen Bildung. Stuttgart 1977, 173–184.

Zur Übertragbarkeit des Beutelsbacher Konsenses auf die Religionspädagogik siehe: KATHARINA MUTH: Orientierung geben und Identitäten stiften? Was der Religionsunterricht in Bezug auf Urteilsbildung leisten kann und sollte und was nicht, in: MIRKA DICKEL/ANKE JOHN/MICHAEL MAY/KATHARINA MUTH/LAURENZ VOLKMAN/MARIO ZIEGLER (Hrsg.), Urteilspraxis und Wertmaßstäbe im Unterricht. Ethik, Englisch, Geschichte, politische Bildung und Religion. Schwalbach 2020, 148–172.

⁶⁴ Vgl. FACHVERBAND PHILOSOPHIE E.V./FACHVERBANDS ETHIK E.V./FORUM FÜR DIDAKTIK DER PHILOSOPHIE UND ETHIK (Hrsg.), Dresdner Konsens für den Philosophie- und Ethikunterricht, in: Zeitschrift für Didaktik der Philosophie und Ethik 38 (2016), 106. Abrufbar unter: https://philosophiedidaktik.files.wordpress.com/2017/03/dresdner_konsens.pdf (16.05.2018).

2.3 Verortung der Studie im religionspädagogischen Diskurs

VEIT-JAKOBUS DIETRICH kommt in einer Metaanalyse unterschiedlicher bildungsadministrativer Vorgaben zu dem Ergebnis, dass vier Versionen des Eingangs der Kompetenzorientierung in die Grundausrichtung des Religionsunterrichts zu unterscheiden sind: 1.) eine allumfassende Kompetenzausrichtung, 2.) Vorbehalte im Hinblick auf den Glauben selbst, 3.) »ein Ausschluss auch von bestimmten religiösen Einstellungen und Haltungen«, sowie zuletzt 4.) »eine explizite Exklusion von Einstellungen und Haltungen überhaupt«.⁶⁵ Diese vier Positionen lassen sich auch im religionspädagogischen Forschungsdiskurs wiederfinden. In Bezug auf die Kompetenz der Urteilsbildung zu kontroversen Fragen im Religionsunterricht kulminieren die vorgebrachten Vorbehalte hinsichtlich des Glaubens, der Einstellungen und Haltungen:

Konstitutiv für die konzeptionelle Ausrichtung des Religionsunterrichts ist – mit Ausnahme von Hamburg – seine konfessionelle Bindung, welche mit einer Bekenntnisorientierung einhergeht. Vor dem Hintergrund des Verhältnisses zwischen Glaube, Bildung und Bekenntnis wurde in der religionspädagogischen Debatte die Übertragbarkeit der Kompetenzorientierung auf den Religionsunterricht intensiv kritisch diskutiert.⁶⁶ Ein wesentliches Argument ist dabei der Verweis auf die unverfügbaren Elemente und Erfahrungen, die Teil des Religionsunterrichts sind und die sich durch den Sinn für Transzendenz oder das »Ergriffensein von dem, was uns unbedingt angeht« – mit PAUL TILICH gesprochen – beschreiben lassen.⁶⁷ Dieses Unverfügbare – so der Einwand – entzieht

⁶⁵ VEIT-JAKOBUS DIETRICH, Wie steht es mit der Kompetenzorientierung in den Lehr- und Bildungsplänen für den Religionsunterricht? Eine Standortbestimmung nach eineinhalb Jahrzehnten Curriculumrevision, in: Zeitschrift für Pädagogik und Theologie 70 (2018), 347–359, hier 351.

⁶⁶ Vgl. exemplarisch: VÖLKER ELSNBAST/DIETLIND FISCHER (Hrsg.), Stellungnahmen und Kommentare zu »Grundlegende Kompetenzen religiöser Bildung«. Münster: 2007; MARTIN ROTHGANGEL/DIETLIND FISCHER (Hrsg.), Standards für religiöse Bildung? Zur Reformdiskussion in Schule und Bildung. Münster 2004; MARTIN ROTHGANGEL, Bildungsstandards Religion. Eine Replik auf verbreitete Kritikpunkte, in: ANDREAS FEINDT/VÖLKER ELSNBAST/PETER SCHREINER/ALBRECHT SCHÖLL (Hrsg.), Kompetenzorientierung im Religionsunterricht. Befunde und Perspektiven. Münster u. a. 2009, 87–97; CLAUB PETER SAJAK, Religionsunterricht kompetenzorientiert: Beiträge aus fachdidaktischer Forschung. Paderborn u. a. 2012; MICHAEL WERMKE (Hrsg.), Bildungsstandards und Religionsunterricht. Perspektiven aus Thüringen. Jena 2005 sowie das Themenheft Kompetenzorientierung im Religionsunterricht – Chancen und Grenzen, in: Zeitschrift für Pädagogik und Theologie 70 (2018), 372–385.

⁶⁷ PAUL TILICH, Gesammelte Werke. Bd. 9, Stuttgart 1967, 94.

sich einer kompetenzterminologischen Logik und einem operationalisier- und messbaren Zugriff, »weil das Subjekt des Religiösen nicht objektiviert werden kann«. ⁶⁸ Mit diesem Argument öffnet sich die Debatte hin zu der Frage nach der generellen Machbarkeit religiösen Lernens, der Verfügbarkeit von Bildung und nach ihren anthropologischen Grundlagen. MARTIN ROTHGANGEL hebt hervor, dass sich die unverfügbaren Dimensionen des Religionsunterrichts grundsätzlich der religionspädagogischen Machbarkeit entzieht – unabhängig davon, ob ein kompetenzorientierter oder ein anderer religionspädagogischer Ansatz verfolgt wird. ⁶⁹ Da die generelle Notwendigkeit religiöser Bildung im Kontext Schule nicht in Frage gestellt wird, ist zu definieren, was auf welche Weise gelehrt und gelernt werden kann.

Als unverfügbar gilt im religionspädagogischen Forschungsdiskurs neben dem Sinn für Transzendenz auch der Glaube an Gott, welcher nicht gelehrt oder gemessen werden kann, sondern aus theologischer Sicht ein unverfügbares Geschenk Gottes ist. ⁷⁰ Dennoch ist »der Glaube an Gott oder eben die Frage nach der Wahrheit der menschlichen Existenz« ⁷¹ nach FRIEDRICH SCHWEITZER das, worum es im Religionsunterricht in erster Linie geht. BERNHARD DRESSLER definiert Glaube als »die individuelle Gewissheit des durch Christus erschlossenen Gottvertrauens«, welche eine unverfügbare Gabe darstellt, während Religion als »Medium der Glaubenskommunikation« durchaus lernbar ist. ⁷² DRESSLER macht deutlich, dass Glaube durch religiöse Kompetenz gedeihen kann, dass er aber

⁶⁸ CAROLIN CRAMER, Chancen und Grenzen der Kompetenzorientierung in Religionsunterricht und Religionslehrerbildung, in: *Zeitschrift für Pädagogik und Theologie* 70 (2018), 372–385, hier 378. Vgl. auch MARTIN ROTHGANGEL, Das Beste am Religionsunterricht? Kompetenzorientierung im anthropologischen Horizont, in: THOMAS SCHLAG/HENRIK SIMOJOKI (Hrsg.), *Mensch – Religion – Bildung. Religionspädagogik in anthropologischen Spannungsfeldern*. Gütersloh 2014, 478–489.

⁶⁹ ROTHGANGEL, *Das Beste am Religionsunterricht?*, 488.

⁷⁰ Vgl. exemplarisch FRIEDRICH SCHWEITZER: Leistungsmessung und Leistungsbewertung, Kompetenzen und Standards: Was ist im Religionsunterricht messbar?, in: *entwurf* 47 (2016), 4–8, hier 4. Sowie GEORG HILGER, Welche Wirkung hat der Religionsunterricht? Evaluation und Leistungsbewertung, in: GEORG HILGER/STEPHAN LEIMGRUBER/HANS-GEORG ZIEBERTZ (Hrsg.), *Religionsdidaktik. Ein Leitfaden für das Studium, Ausbildung und Beruf*. München ⁶2003, 282–290, hier 285 f. Sowie in einer Metaanalyse bildungsadministrativer Vorgaben DIETERICH: Kompetenzorientierung, 350 f.

⁷¹ FRIEDRICH SCHWEITZER, Religiöse Bildung ohne Ethik? Zur ethischen Dimension des Religionsunterrichts, in: RUDOLF ENGLERT/HELGA KOHLER-SPIEGEL/ELISABETH NAURATH/BERND SCHRÖDER/FRIEDRICH SCHWEITZER (Hrsg.), *Ethisches Lernen*. Neukirchen-Vluyn 2015 [= *Jahrbuch der Religionspädagogik* 31], 13–23, hier 23.

⁷² BERNHARD DRESSLER: Religionsunterricht – mehr Kompetenzorientierung? In: ANDREAS FEINDT/VOLKER ELSENBAST/PETER SCHREINER/ALBRECHT SCHÖLL (Hrsg.), *Kompetenzorientierung im Religionsunterricht. Befunde und Perspektiven*. Münster u. a. 2009, 23–37, hier 28.

dennoch nicht »erzeugbar« ist.⁷³ Mit Blick auf Urteilsfähigkeit in ethischen und religiösen Fragestellungen ist daher zu diskutieren, welche Rolle sowohl der Glaube als auch Einstellungen und Haltungen in diesen Fragen einnimmt und wie sich die Spannung zwischen Erlern- und Unverfügbarem in Bezug auf Urteilskompetenz beschreiben lässt.

Im religionspädagogischen Diskurs ist – nicht zuletzt aufgrund der konfessionellen Bindung – neben den in der Kompetenzorientierung verankerten Begriffen der Einstellungen und Haltungen der Begriff der Werte zentral. Werte können definiert werden als wünschenswerte, erforderliche und erstrebenswerte Zustände.⁷⁴ Einstellungen sind im Gegensatz zu Werten und Haltungen – die sich abstrakt auf ganze Objektbereiche beziehen – auf bestimmte Objekte gerichtet und können definiert werden als »zusammenfassende Bewertung eines Gegenstands«, welche sich in Ablehnung oder Zuneigung zu einem bestimmten Objekt äußert.⁷⁵ Einstellungen, Haltungen und Werten ist gemein, dass sie kognitive, affektive und pragmatische (im Sinne von »handlungssteuernde«) Aspekte besitzen.⁷⁶ In der religionspädagogischen Debatte werden die Begriffe weitestgehend synonym gebraucht.

⁷³ Vgl. a. a. O., 28 f.

⁷⁴ Die Theologische Realenzyklopädie definiert: »Als ›Werte‹ lassen sich diejenigen Gegebenheiten irgendwelcher Art bezeichnen, die ein Bedürfnis oder ein Interesse befriedigen, ein Lustgefühl hervorrufen, Anerkennung verdienen, wünschenswert, erforderlich oder erstrebenswert erscheinen.« WOLFHART HENCKMANN/ HARTMUT KREß, Art. Wert, in: Theologische Realenzyklopädie Bd. 35, 2003, 648–657, hier 648.

⁷⁵ GERD BOHNER, Einstellungen, in: WOLFGANG STROEBE/KLAUS JONAS/MILES HEWSTONE (Hrsg.), Sozialpsychologie. Eine Einführung. Berlin u. a. 2002, 266–314, hier 267.

⁷⁶ Vgl. OLIVER REIS, Zwischen studentischer Diversität und fachlicher Homogenisierung. Die wertbezogene Hochschuldidaktik als Vermittlerin im Habituskonflikt, in: SYLVIA HEUCHEMER/FRIEDERIKE SILLER/TIMO VAN TREECK (Hrsg.), Hochschuldidaktik forscht zu Vielfalt und Offenheit. Profilbildung und Wertefragen in der Hochschulentwicklung I. Cologne Open Science: 2018. Abrufbar unter: https://cos.bibl.th-koeln.de/frontdoor/deliver/index/docId/802/file/FIHB_Band_2_final_web.pdf (23.1.2019), 19–49, hier 24.

Einstellungen und Werthaltungen sind an dieser Stelle abzugrenzen von dem derzeit intensiv erforschten Einfluss epistemologischer Überzeugungen der Lehrenden und Lernenden zu einem Wissensgebiet auf den Lernerfolg. »Epistemologische Überzeugungen bezeichnen individuelle subjektive Ansichten, Auffassungen und Theorien über die Genese, Ontologie, Bedeutung, Rechtfertigung und Gültigkeit von Wissen in den Wissenschaften. Sie spiegeln damit Grundauffassungen und Einstellungen gegenüber Wissenschaften wider« (vgl. BURKHARD PRIEMER, Deutschsprachige Verfahren der Erfassung von epistemologischen Überzeugungen, in: Zeitschrift für Didaktik der Naturwissenschaften 12 [2006]), 159–175, hier 160). Es handelt sich also um spezifische Überzeugungen, nämlich solche zum Wissen und zum Wissenserwerb. Diese sind wiederum abzugrenzen von subjektiven Theorien. Unter subjektiven Theorien versteht man Kognitionen des Alltagsmenschen, »die sich inhaltlich auf

Welche Positionen sind also hinsichtlich der Ausbildung spezifischer Einstellungen respektive Werte im Religionsunterricht zu benennen? Die Autorinnen und Autoren des vom Comenius-Institut in Auftrag gegebenen Modells religiöser Kompetenz beispielsweise schreiben auf die Klieme-Expertise und der dort vertretenen WEINERT'schen Definition von Kompetenzen Bezug nehmend, dass es ein »ausdrückliches Ziel« des Religionsunterrichts sei, Werthaltungen auszubilden.⁷⁷ Sie konstatieren jedoch eine schwere Vereinbarkeit dieses Ziels mit dem Anspruch der Kompetenzorientierung, alle Learning-Outcomes operationalisieren, messen und bewerten zu können:

»In dem Kompetenzmodell von WEINERT geht es nicht nur um bereichsspezifische Wissensinhalte, Fähigkeiten und Fertigkeiten, sondern auch um deren Verknüpfung mit Einstellungen, Werthaltungen und Motivationen. Ein Beitrag zur Ausbildung von Werthaltungen wie z.B. Verantwortungsbereitschaft, Solidarität, Respekt und Sensibilität ist ausdrückliches Ziel des Religionsunterrichts. Es ist jedoch offen, wie derartige Einstellungen und Werthaltungen festgestellt werden können, ohne Schüler/innen zu Bekenntnisakten zu nötigen.«⁷⁸

Eine Gegenposition nimmt BERNHARD DRESSLER ein. Er widerspricht der Annahme, dass Schule überhaupt die Aufgabe haben sollte, *spezifische* Haltungen auszubilden. Er schreibt:

»Die Lehrbarkeit von Moralität (im Unterschied zur ethischen Urteilsfähigkeit) ist weder operationalisierbar noch wünschbar, weil das bildende Selbstverständnis der Lernenden zu ihren Lerngegenständen der Bedingung unterliegt, keine besondere Werte- oder Haltungserziehung abzuspalten. Moralerziehung liefe darauf hinaus, mit pädagogischen Begründungen jene Optionen wieder auszuschalten, die im Bil-

seine Selbst- und Weltsicht beziehen und eine vergleichbare Struktur und Funktion aufweisen wie »objektive« Theorien für den/die Wissenschaftler/in« (URSULA CHRISTMANN/NORBERT GROEBEN/MARGRIT SCHREIER, Subjektive Theorien – Rekonstruktion und Dialog-Konsens, in: SPIEL 18 [1999], 138–154, hier 138 f.).

⁷⁷ DIETLIND FISCHER/VOLKER ELSNBAST (Redaktion), Grundlegende Kompetenzen religiöser Bildung. Zur Entwicklung des evangelischen Religionsunterrichts durch Bildungsstandards für den Abschluss der Sekundarstufe I. Münster 2006, 16.

Von dieser Position geht auch BRITTA KLOSE aus, die mit einer empirischen Studie über die diagnostischen Wahrnehmungskompetenzen von Religionslehrkräften die Erkenntnis unterstreicht, dass eine passgenaue Wahrnehmung der Werthaltungen der Schülerinnen und Schüler durch die Religionslehrkraft einen positiven Einfluss auf die Zufriedenheit der Schülerinnen und Schüler mit dem Religionsunterricht sowie auf die Möglichkeiten der Schülerorientierung hat (vgl. DIES., Diagnostische Wahrnehmungskompetenzen von ReligionslehrerInnen. Stuttgart 2014, insbesondere 194 sowie 214).

⁷⁸ Ebd.

dungsprozess freizugeben sind. Die Schule hat die Möglichkeit der begründeten und begründenden Wertung, nicht aber Werte zu lehren. Grundlegend ist die Schule auch dann, wenn dem in ihr gestaltenden Leben eine erzieherische Funktion und Intention nicht abgesprochen werden kann, kein Ort, an dem Identitäten gestiftet werden sollen.«⁷⁹

Werteerziehung, so DRESSLER, stehe den Schulen grundsätzlich nicht zu.⁸⁰ Bezugnehmend auf DRESSLERS Position schreibt FRIEDRICH SCHWEITZER:

»Was spräche dafür, in der Schule von vornherein auf jede Beeinflussung von Haltungen verzichten zu wollen – beispielsweise auch einer Haltung wie Intoleranz und Diskriminierung? Würde sich eine solche Schule nicht selbst ad absurdum führen? Und folgt aus dem Hinweis, dass Schule nicht einfach Werte ›lehren‹ kann (auch das ist ja richtig), auch schon, dass sie überhaupt keinerlei Möglichkeit hat und haben soll, zur Wertebildung wenigstens das Ihre beizutragen?«⁸¹

DRESSLER und SCHWEITZER stimmen in der Auffassung überein, dass man Werte weder kausal vermitteln noch operationalisieren kann und dass die Lernenden immer als mündige Subjekte angesprochen sein müssen.⁸² Während SCHWEITZER es dennoch für eine wichtige Aufgabe des Religionsunterrichts hält, spezifische Werte anzubahnen, lehnt DRESSLER Werteerziehung im Religionsunterricht ab. Da Werte nicht einfach vermittelt werden können, ist es DRESSLER zu folge sowohl töricht als auch eine Überforderung, von der Schule zu erwarten, dass sie »als gesellschaftliche[r] Reparaturbetrieb[] für anderswo entstandene Schäden«

⁷⁹ BERNHARD DRESSLER, *Unterscheidungen. Religion und Bildung*. Leipzig 2006, 186 f. Bei der hiesigen Ausführung über Moralität nimmt Dressler Bezug auf DIETRICH BENNER, *Allgemeine Pädagogik. Eine systematisch-problemgeschichtliche Einführung in die Grundstruktur pädagogischen Denkens und Handelns*. Weinheim/Basel ⁸2015, 256–258.

⁸⁰ Vgl. DRESSLER, *Unterscheidungen*, 187.

⁸¹ SCHWEITZER, *Religiöse Bildung ohne Ethik?*, 21.

⁸² DRESSLER führt zur Operationalisierung von Werten aus: »Moralische Dispositionen und Präferenzen entstehen in sozialen Handlungszusammenhängen und hängen an den darin gemachten Erfahrungen. Nur durch Bildung zur Religion (also aus einer bestimmten Welt- und Selbstdeutung heraus) oder durch Bildung zur ethischen Urteilsfähigkeit kann auch sozial erwünschtes Handeln in den Blick geraten – aber im Sinne freier, und eben darum auch stabiler Selbstorientierung. Emotionen, Haltungen und Einstellungen, die im Zuge des Bildungsprozesses entstehen, stehen weder den Pädagogen noch den Bildungspolitikern zur Disposition – und zwar de facto weil das pädagogische Handeln hier auf eine Grenze der Operationalisierbarkeit stößt, aber auch aus Gründen der Würde der zu Erziehenden« (DRESSLER, *Unterscheidungen*, 189).

Vgl. zur Debatte über die Lehrbarkeit von Werten auch: MUTH, *Orientierung geben*, 148–172.

fungieren könne, wie es in der Forderung, der Religionsunterricht solle Werteerziehung leisten, mitschwingt.⁸³ Zudem geht es DRESSLER darum, die Funktionalisierung von religiöser Bildung entschieden abzulehnen.

Gleichzeitig fordert DRESSLER, dass der Religionsunterricht über die ethische Urteilsbildung hinausgeht und nicht auf Metaphysik oder Ethik reduziert wird.⁸⁴ Er plädiert für eine Gewichtsverschiebung im Religionsunterricht weg von Fachwissen zugunsten elementaren religiösen Lernens (im Sinne von Performanz religiöser Praxis; die Spuren der Liebe Gottes in der Welt sichtbar machen).⁸⁵ Es bleibt jedoch fraglich, wohin eine Teilhabe an religiöser Praxis und das Wahrnehmen der »Spuren der Liebe Gottes« führen sollen, wenn nicht zu einer Veränderung von Haltungen, Einstellungen und Weltansichten – also Wertebildung.

An die Debatte um die Relevanz von Einstellungen, Haltungen und Werten im Lernprozess knüpft unmittelbar die Frage nach der Prüfbarkeit und Bewertung subjektiver Sichtweisen an. Eine Position, die als weitestgehend anerkannt in dieser Frage gelten kann, bringt FRIEDRICH SCHWEITZER in seinem Artikel *Leistungsmessung und Leistungsbewertung, Kompetenzen und Standards: Was ist im Religionsunterricht messbar?* pointiert zum Ausdruck:

»Bewertet werden können Wissen und Verstehen, vor allem im Sinn der Wahrnehmungs- und Urteilsfähigkeit, aus unterschiedlichen Gründen aber nicht Glaube, Handlungskompetenz sowie Einstellungen.«⁸⁶

Bemerkenswert ist, dass SCHWEITZER im Zitat Einstellungen argumentatorisch von Urteilsfähigkeit abgegrenzt, die er mit den Fähigkeiten Wissen und Verstehen verbindet. Dabei bleibt die Frage unbeantwortet, wie Einstellungen aus der Urteilsbildung hinsichtlich Diagnose- und Bewertungsfragen ausgeklammert werden können. Dass in den Abituraufgabenstellungen des Faches Evangelische Religionslehre Urteilsbildungsaufgaben in Bezug auf objektiv nicht entscheidbare Fragen sowohl bildungsadministrativ vorgesehen als auch in der Praxis zu finden sind, macht deutlich, dass prinzipiell nicht in Frage gestellt wird, dass man

⁸³ Vgl. DRESSLER, Unterscheidungen, 186.

⁸⁴ Vgl. BERNHARD DRESSLER, Religionsunterricht als Werteerziehung? Eine Problemanzeige, in: Zeitschrift für Evangelische Ethik 46 (2002), 256–269, hier u. a. 264.

⁸⁵ Vgl. a. a. O., 264 f. DRESSLER räumt auch die reflexive Distanz ein, welche er als notwendig für Lernprozesse hält.

⁸⁶ SCHWEITZER, Leistungsmessung, 7. Die dem Zitat zugrunde gelegte Dichotomie zwischen Urteilsfähigkeit und Einstellungen darf nicht zu einer vereinfachten Trennung beider Bereiche verleiten, da normative Wertungen gefragt sind, bei denen Einstellungen im Urteilsbildungsprozesse eine wesentliche Rolle einnehmen. Richtig bleibt dennoch, dass die Einstellung an sich nicht bewertet werden darf. Diesem Faktum muss auch die Aufgabenstellung Rechnung tragen.

Urteilsbildung prüfen und bewerten kann. In Bezug auf die Einstellungen führt SCHWEITZER nochmals aus:

»Selbstverständlich vorausgesetzt wird dabei [bei der Notengebung im Religionsunterricht; Anm. K.M.], dass sich diese Bewertung weder auf den Glauben, noch auf persönliche Einstellungen beziehen kann, sondern eben nur auf das, was im Unterricht tatsächlich gelernt werden kann. Im Blick auf Einstellungen gilt also [...]: Gemessen werden können sie, aber benotet werden dürfen sie nicht.«⁸⁷

GABRIELE OBST positioniert sich ähnlich wie SCHWEITZER in diesem Punkt. Sie schreibt bezüglich der Grenze zwischen Prüfbarem und Nicht-Prüfbarem, das nur das im Unterricht Erlernbare überprüft werden sollte und grenzt Identitätsentwicklung und Einstellungen davon ab:

»Ein schlichtes aber wirkungsvolles Kriterium für eine solche Grenze [des Bewertbaren, Anm. KM] könnte sein, dass Kompetenzen in dem beschriebenen begrenzten Sinn sich auf das beziehen, was durch organisierten Unterricht erlernbar ist, nicht aber auf das, was in den Bereich persönlicher Identitätsentwicklung, individueller Überzeugung und Einstellungen sowie religiöser Glaubens- und Lebenspraxis hineinreicht, ebenso wenig natürlich wie in die entsprechenden Optionen für atheistische oder agnostische Positionen oder in andere religiöse Präferenzen.«⁸⁸

Sowohl SCHWEITZER als auch OBST machen durch ihre Äußerungen deutlich, dass Einstellungen ihrer Meinung nach nicht bewertet werden sollen, da sie nicht zu dem Bereich gehören, der »im Unterricht tatsächlich gelernt werden kann«. Diese Position baut auf einer Differenzierung zwischen erlernbarem Wissen und zu glaubendem Wissen auf, welche REIS und SCHWARZKOPF problematisieren, da man mit dem Ausschluss des zu glaubenden Wissens aus Lernzielen und Bewertungen auch für diesen Bereich keine Lernentwicklung formulieren kann.⁸⁹ Aus ihrer Sicht stellen die Einstellung zum Lerngegenstand jedoch einen wesentlichen Bestandteil in der Definition des Learning-Outcomes dar. Sie schreiben:

»[W]er die persönliche Beziehung zum Inhalt der Bewertung entzieht, um die Person zu schützen, kommt nicht aus der Polarisierung von faktischen Inhalten und zu glaubenden Inhalten heraus. Wenn wir dagegen den Lernbereich unter Lernzielen

⁸⁷ A. a. O., 6.

⁸⁸ GABRIELE OBST, Kompetenzorientiertes Lehren und Lernen im Religionsunterricht. Göttingen ⁴2015, 254 f.

⁸⁹ Vgl. REIS/SCHWARZKOPF, Diagnose religiöser Lernprozesse, 46.

genau von einem Bereich ohne Lernziele abgrenzen, dann fordern wir für den Bereich unter Lernzielen solche Ziele, die die Beteiligung einer Person erfordern und dass eben auch zu Fragen, die nicht mehr objektiv entscheidbar sind.«⁹⁰

REIS und SCHWARZKOPF fordern also, dass die Beteiligung der Person nicht aus den Lernzielen ausgeschlossen, sondern bewusst und explizit in diese aufgenommen wird. Damit widersprechen sie bewusst der Position, dass Mehrperspektivität Mehrdeutigkeit bedeutet und dass diese nur persönlich verantwortet und persönliche Positionierungen nicht bewertet werden können.⁹¹ In ähnlicher Weise argumentiert HERBERT A. ZWERGEL, wenn er betont, dass »gerade die sich einem operationalisierenden Zugriff entziehenden Entwicklungsmomente einer Durchdringung bedürfen, die der Komplexität der zu fördernden Fähigkeiten, Bereitschaften und Haltungen gerecht wird. Damit diese nicht einem diffusen Einschätzungsraum überantwortet werden, in dem man »nichts machen« könne, weil nichts überprüfbar sei, muss die pädagogisch-didaktische Durchdringung umso gründlicher ansetzen.«⁹² REIS und SCHWARZKOPF schlagen in einem von ihnen vorgelegten Diagnosemodell vor, für »die nicht mehr objektiv entscheidbaren« Fragen einen didaktischen Rahmen zu entwickeln, den sie als *Meta-Struktur*⁹³ bezeichnen und

»[...] der die Offenheit der Denkmöglichkeiten als Instrument zugleich so strukturiert, dass die Tätigkeit an Objekten an einem Maßstab gemessen werden kann. Wenn nicht die Freiheit der Wahl aufgehoben werden soll, dann muss diese Struktur den Lerngegenstand mehrperspektivisch fassen, so dass der Maßstab hier die Qualität der Tätigkeit sein kann, die wesentlich durch die Qualität des Zugriffs auf die Gesamtstruktur der Antwortmöglichkeiten erfolgt.«⁹⁴

Das von REIS und SCHWARZKOPF vorgeschlagene Diagnosemodell zielt darauf ab, eine Norm zu benennen, die Normabweichungen sichtbar werden lassen. Diese Norm wird von den Learning-Outcomes her entwickelt, welche sich aus den Kernlehrplänen ableiten.⁹⁵ Die Prüfbarkeit einer Kompetenzerwartung ist somit von der Formulierung des Learning-Outcomes abhängig, was REIS und

⁹⁰ A. a. O.

⁹¹ Vgl. a. a. O., 73.

⁹² HERBERT A. ZWERGEL, Kompetenzen und Grenzen ihrer Standardisierung – Umgang mit ethischen Entscheidungssituationen und existenziellen Grenzsituationen am Beispiel religiösen Lernens, in: VÖLKER FREDERKING (Hrsg.), Schwer messbare Kompetenzen. Herausforderungen für die empirische Fachdidaktik. Baltmannsweiler 2008, 85–114, hier 87.

⁹³ Die Schreibweise entspricht der von SCHWARZKOPF und REIS genutzten Schreibweise.

⁹⁴ REIS/SCHWARZKOPF, Diagnose religiöser Lernprozesse, 67.

⁹⁵ A. a. O., 105.

SCHWARZKOPF an folgendem Beispiel illustrieren: Das Learning-Outcome, »Die Schülerinnen erkennen, dass die biblischen Gebote dem Zusammenleben dienen«, ist REIS und SCHWARZKOPF zufolge »schlichtweg falsch gebaut«, da es die Würde der Persönlichkeit nicht achtet.⁹⁶ Als ein ohne Abstriche diagnostizierbares Learning-Outcome schlagen sie hingegen eine offenere Formulierung vor: »Die Schülerinnen bewerten die Güte biblischer Gebote, das Zusammenleben zu [sic!] humaner zu machen.«⁹⁷ Von Learning-Outcomes ausgehend, die Mehrperspektivität ermöglichen, plädieren REIS und SCHWARZKOPF für eine modellgeleitete Diagnostik anhand von Niveaustufendifferenzierungen (zu Niveaustufendifferenzierung siehe Kap. 3.3). Somit soll auch die von REIS und SCHWARZKOPF vorgeschlagene MetaStruktur die Mehrperspektivität einer Fragestellung anhand konkreter Positionen des Fachdiskurses abbilden und dadurch einen modellbezogenen Blick auf die zu beurteilende Frage erzeugen. Als Qualitäts- und damit Bewertungsmaßstab wird der Umgang mit der Fragestellung – also der Zugriff auf die Gesamtstruktur – benannt, nicht aber die Wahl einer Option selbst.

Anhand der weiteren von REIS und SCHWARZKOPF beispielhaft formulierten Learning-Outcomes wird deutlich, dass die Kohärenz der eigenen Argumentation, also die innere Schlüssigkeit der hergestellten Zusammenhänge, aus ihrer Sicht wesentlich für die Beurteilung des Umgangs mit der Fragestellung ist und somit für die MetaStruktur eine Systematisierung bietet. Ihrem Anspruch gerecht zu werden, »sowohl die taxonomische Komplexität von Lernzielen als auch die Verbindlichkeit in der Auseinandersetzung der Person mit den Inhalten zu erhöhen«, setzt umfangreiche Kenntnis des Fachdiskurses zur Fragestellung und der einzelnen Positionen bei der Konzeption der MetaStruktur voraus.⁹⁸ REIS und SCHWARZKOPF sind davon überzeugt, dass es möglich ist, trotz Wahrung der Wahlfreiheit den Schülerinnen und Schülern eine Abweichung vom Normalverfahren deutlich zu machen und dabei zu helfen, »die angemesseneren [im Sinne von konsistenteren, Anm. KM] Entscheidungen zu treffen.«⁹⁹ Diese Sicht geht davon aus, dass es auch in Bezug auf objektiv nicht mehr zu entscheidende Fragen eine »quasi-objektiv vorausgesetzte normierte Handlungsweise gibt, die angemessene Lösungen von unangemessenen Lösungen unterscheidet«, wobei nicht die Mehrzahl möglicher Lösungen in Frage gestellt wird. Es wird jedoch betont, dass es entscheidende Schnittstellen gibt, an denen man zu unangemessenen Lösungen kommen kann, beispielsweise, weil sie komplexe Deutungen verhindern. REIS und SCHWARZKOPF sehen es als eine wesentliche Aufgabe eines kompetenzorientierten Religionsunterrichts, diese Normabweichungen auch für den Bereich der unentscheidbaren Fragen methodisch angelegt defi-

⁹⁶ A. a. O., 44.

⁹⁷ Ebd.

⁹⁸ A. a. O., 68.

⁹⁹ Ebd.

nieren zu können, um aus der Struktur der Tätigkeit pädagogische Maßnahmen ableiten zu können.¹⁰⁰

REIS und SCHWARZKOPF stimmen folglich mit der oben exemplarisch durch SCHWEITZER und OBST zitierten Position überein, dass spezifische Einstellungen, Haltungen und Werte im Sinne einer normativ gesetzten Wahl nicht zu bewerten sind. Dies würde der Mehrperspektivität und der Komplexität von Urteilsbildungsaufgaben nicht gerecht. Eben in diesem Argument stimmen sie sogar mit DRESSLER überein, der einen Verzicht auf die Ausbildung spezifischer Werte im Unterricht fordert, dem SCHWEITZER zustimmen würde. Ihnen greift jedoch die Schlussfolgerung zu kurz, dass man die Ausbildung von normativ-gesetzten Einstellungen zu einem Lerngegenstand deshalb nicht als Outcome definieren und bewerten kann. Im Gegenteil fordern sie klar definierte Lernziele und eine klare Beschreibung des Kompetenzerwerbs für dieses Feld, welche messbar sein und Rückschlüsse auf pädagogische Fördermaßnahmen bieten sollten. Der Ansatz von REIS und SCHWARZKOPF bietet die Chance, Einstellungen und Haltungen in komplexen, kompetenzorientierten Anforderungssituationen, wie sie eine Urteilsbildung darstellt, nicht künstlich aus dem Bewertungsprozess auszuklammern, wie es in der religionspädagogischen Debatte durch die Differenzierung zwischen Urteilsbildung und Einstellungen oder Werten häufig geschieht, sondern sie aktiv mit zu bedenken und hinsichtlich ihrer inhaltlichen Kohärenz zu bewerten.

¹⁰⁰ Ebd.

2.4 Forschungsanliegen und Forschungsfragen

Das Erkenntnisinteresse der Studie ist in erster Linie deskriptiv: Es soll anhand schriftlicher Abituraufgabenstellungen aus dem Fach Evangelische Religionslehre und durch die Analyse der dazugehörigen Erwartungshorizonte untersucht werden, wie mit Einstellungen und Haltungen im Kontext von Urteilsbildungsaufgaben umgegangen wird. Ziel der Studie ist eine Reflexion des den Aufgabenstellungen und Bewertungsvorgaben zugrunde liegenden Konstrukts von Urteilsfähigkeit vor dem Hintergrund des aktuellen Forschungsdiskurses. Ein besonderer Fokus liegt auf der Fragestellung, wie mit der Herausforderung der Bewertung von Urteilsfähigkeit in kontroversen ethischen und religiösen Fragen umgegangen wird, die durch den Anspruch entsteht, die Offenheit der Urteilsbildung zu wahren und dennoch zu allgemeinverbindlichen Bewertungskriterien zu kommen. Als Datenkorpus dienen die Abituraufgabenstellungen und Erwartungshorizonte im Fach Evangelische Religionslehre der exemplarisch ausgewählten Bundesländer Bayern, Niedersachsen und Thüringen aus den Jahren 2014 bis 2019, welche in einer inhaltlich strukturierenden qualitativen Inhaltsanalyse analysiert werden.

Durch die Inhaltsanalyse wird zum einen eruiert, *wie* Urteilsfähigkeit derzeit im schriftlichen Abitur des Faches Evangelische Religionslehre geprüft wird. Zentral sind dabei folgende Fragestellungen:

- *Wie häufig wird eine eigenständige Urteilsfähigkeit von den Prüflingen verlangt?*
- *Anhand welcher Fragestellungen wird Urteilsfähigkeit im schriftlichen Abitur des Faches Evangelische Religionslehre geprüft?*
- *Welche Rolle spielt eine christliche Weltansicht in Urteilsbildungsaufgaben im Fach Evangelische Religionslehre?*

Darüber hinaus wird in Anknüpfung an die im Vorangegangenen skizzierte religionspädagogische Debatte um die Bewertbarkeit von Einstellungen, Haltungen und Wertungen nach den Bewertungskriterien für Urteilsfähigkeit gefragt. Fokussiert werden folgende Fragestellungen:

- *Welche Qualitätsmerkmale benennen die Erwartungshorizonte der schriftlichen Abituraufgabenstellungen für Urteilsfähigkeit in ethischen und religiösen Fragen?*
- *Geben die Erwartungshorizonte explizite Hinweise auf den Umgang mit Einstellungen und Haltungen in Bezug auf Urteilsbildungsaufgaben?*
- *Welche Unterschiede lassen sich zwischen den exemplarisch untersuchten Bundesländern Bayern, Niedersachsen und Thüringen hinsichtlich ihrer Anforderungen an und ihrer Bewertungsmaßstäbe für Urteilsfähigkeit ausmachen?*

3 Diagnostische Rahmenbedingungen von Urteilskompetenz in schriftlichen Abiturprüfungen – Forschungshintergrund

Als Forschungshintergrund der Studie und gleichzeitig zur besseren Einordnung der Ergebnisse sind drei Aspekte zentral: Zum einen die Postulate von Vergleichbarkeit und Transparenz, welche durch zentrale Abschlussprüfungen erreicht werden sollen, die sich an Einheitlichen Prüfungsanforderungen (EPA) orientieren. Sie werden in Kapitel 3.1 diskutiert. Dabei werden pragmatische Fragen zur Erstellung zentraler Abituraufgabenstellungen beantwortet, die im Kontext der Analyse und Bewertung der Abituraufgabenstellungen aufkommen.

Zum anderen gilt es als Vorüberlegung für die Analyse, die bundesweit verbindlichen bildungsadministrativen Vorgaben für Urteilskompetenz im Fach Evangelische Religionslehre zu reflektieren, die in den EPA festgeschrieben sind (vgl. Kapitel 3.2). An die Betrachtung der staatlichen Vorgaben schließt sich in einem Exkurs ein Blick in die kirchlichen Verlautbarungen an, da der Religionsunterricht laut Verfassung inhaltlich durch die Kirche mitbestimmt wird. Es wird daher der Frage nachgegangen, welche Stellung die kirchlichen Denkschriften zum Religionsunterricht der Kompetenz Urteilsfähigkeit beimesen.

Als dritter Aspekt sind diagnostische Grundlagen und Besonderheiten in Bezug auf Urteilsfähigkeit von Relevanz für die Forschungsfrage, da sich Bewertungskriterien an diesen ausrichten müssen. In Kapitel 3.3 wird daher auf verschiedene Formen von Kompetenzmodellen, auf Bezugsnormen, Gütekriterien, Bewertungsansätze und auf das Medium der Schriftlichkeit, in dem die untersuchten Abituraufgabenstellungen zu lösen sind, Bezug genommen.

3.1 Vergleichbarkeit zwischen zentralen Abschlussprüfungen

Die schriftlichen Abituraufgabenstellungen für Evangelische Religionslehre werden – ebenso wie die Aufgabenstellungen der anderen Fächer auch – von einer schulexternen Instanz, in der Regel den Kultusministerien der Länder, für jedes Bundesland zentral herausgegeben. Die Einführung zentraler Abschluss-

prüfungen für den mittleren und den gymnasialen Schulabschluss vollzog sich im Rahmen bildungspolitischer Reformen seit der Jahrtausendwende – mit Ausnahme von Rheinland-Pfalz – in der gesamten Bundesrepublik.¹ In den in der vorliegenden Studie untersuchten Bundesländern Bayern, Thüringen und Niedersachsen haben zentrale Prüfungsaufgaben eine unterschiedlich lange Tradition. Bayern führte 1946 als eines der ersten Bundesländer zentrale Abschlussprüfungen ein. Thüringen knüpft 1990 an die Tradition der DDR an, die seit 1959 zentrale Abschlussprüfungen durchführte. In Niedersachsen werden erst seit 2006 zentrale Abituraufgaben herausgegeben.²

Als Bezugsrahmen für die Erstellung der Aufgaben dienen die Curricula der jeweiligen Bundesländer sowie die bundesweit gültigen EPA des jeweiligen Faches. Mit der Einführung des Zentralabiturs wurde die Hoffnung verbunden, durch bundesweit einheitliche Prüfungsvorgaben sowie landesweit einheitliche Prüfungsaufgaben und einheitlich geltende Erwartungshorizonte die Vergleichbarkeit und Transparenz der Vergabe von Schulabschlüssen zu erhöhen, welche durch die föderalistische Schulorganisation und der damit verbundenen Vielfalt an Schulformen und Oberstufenbestimmungen stark in die Kritik geraten ist.³ Die Vergleichbarkeit zwischen den Bundesländern ist vor allem deshalb entscheidend, da es eine bundesweit gültige Hochschulzugangsbedingung, den sogenannten *Numerus Clausus*, gibt und ungleiche Bedingungen daher Ungerechtigkeit in der Vergabe von Studienplätzen bedingen. Zudem proklamieren Befürworter der einheitlichen Abschlussprüfungen, dass diese zu einer Qualitätssicherung des deutschen Bildungssystems beitragen würden und sich ihre normierende Wirkung bis auf den vorgelagerten Unterricht und damit folglich

¹ Vgl. STEPHAN OTTO, »Interpretieren Sie das Gedicht...« Aufgaben in zentralen Abschlussprüfungen im Fach Deutsch, in: STEFAN KELLER/CHRISTIAN REINTJES (Hrsg.), Aufgaben als Schlüssel zur Kompetenz. Didaktische Herausforderungen, wissenschaftliche Zugänge und empirische Befunde. Münster/New York 2016, 142–151, hier 142. In landeszentralen Prüfungen ergänzen sich in der Regel zentrale (bspw. einheitliche Prüfungsaufgaben) und dezentrale Momente (bspw. schulinterne mündliche Prüfungen und schulinterne Korrekturen) (vgl. E. DOMINIQUE KLEIN/MIRKO KRÜGER/SVENJA M. KÜHN/ISABELL VAN ACKEREN, Wirkungen zentraler Abschlussprüfungen im Mehrebenensystem Schule. Eine Zwischenbilanz internationaler und nationaler Befunde und Forschungsdesiderate, in: Zeitschrift für Erziehungswissenschaft 17 [2014], 7–33, hier 23f.).

² Vgl. SVENJA MAREIKE KÜHN, Steuerung und Innovation durch Abschlussprüfungen? Wiesbaden 2010, 42.

³ Vgl. a. a. O., 44.

Zur überwältigenden Vielfalt der Schulformen in Deutschland siehe: BERTELSMANN STIFTUNG (Hrsg.), Chancenspiegel – eine Zwischenbilanz. Zur Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme seit 2002. Gütersloh 2017.

auch auf die Qualitätsentwicklung und Innovation in Schulen auswirkt.⁴ Kritiker zentraler Prüfungsaufgaben befürchten hingegen, dass durch diese nur ein eingeschränktes Aufgabenspektrum zum Einsatz kommt, wodurch Kanonisierungen sich wiederholender Fachinhalte zustande kommen. Die normierende Wirkung der Prüfungsaufgaben auf den vorgelagerten Unterricht wird sowohl von Befürwortern als auch von Kritikern nicht in Frage gestellt. Man spricht vom sogenannten *Washback-Effekt*, der in der Literatur als Annahme formuliert wird⁵: So schreiben beispielsweise ANDREAS BÜCHTER und ANDREAS PALLACK, welche die impliziten Standardsetzungen durch zentrale Prüfungen empirisch anhand des Mathematikunterrichts untersucht haben, dass unterrichtliche Entscheidungen durch zentrale Abschlussprüfungen bedingt werden.⁶

»Zugleich«, so BÜCHTER und PALLACK, »bilden die Lehrkräfte gemeinsam mit den Schülerinnen und Schülern bei der Vorbereitung auf die zentralen Prüfungen ein Team, da sie anders als beim Unterricht im Vorfeld selbst gestellter Klassenarbeiten nicht wissen, welche Aufgaben ›dran kommen‹. Lehrkräfte orientieren sich dabei stark an den Prüfungsarbeiten der vergangenen Jahre, sodass die zentralen Prüfungen erhebliche inhaltliche Auswirkungen auf den Unterricht haben.«⁷

Der Einfluss zentraler Prüfungsaufgaben auf den vorangehenden Unterricht unterstreicht die Bedeutung ihrer Analyse und Reflexion, wenngleich SVENJA MAREIKE KÜHN darauf hinweist, dass dieser Effekt empirisch noch nicht hinreichend belegt ist.⁸ Breit diskutiert und stark problematisiert wird unter dem

⁴ Vgl. KÜHN, Steuerung, 44.

⁵ Zum *backwash-Effekt* vgl. LIYING CHENG/ANDY CURTIS, Test impact and washback. Implications for teaching and learning, in: CHRISTINE COOMBE/BARRY O'SULLIVAN/PETER DAVIDSON/STEVEN STOYNOFF (eds.), Cambridge Guide to Second Language Assessment. Cambridge 2012, 89–95.

⁶ Vgl. ANDREAS BÜCHTER/ANDREAS PALLACK, Methodische Überlegungen und empirische Analysen zur impliziten Standardsetzung durch zentrale Prüfungen, in: Journal für Mathematik-Didaktik 33 (2012), 59–85 (2012), hier 81. Vgl. in ähnlicher Weise bspw. JULIANE KÖSTER, Lern- und Leistungsaufgaben im Deutschunterricht, in: Deutschunterricht 61 (2008), 4–10 oder DISDORN-LIESEN, Vergleichbarkeit, 7 f. sowie STEPHAN OTTO/SVENJA MAREIKE KÜHN, Zwischen Tradition und Innovation – Eine Analyse zentraler Prüfungsaufgaben im Fach Deutsch im Nationalen Vergleich, in: Beziehungen in Unterricht und Schule 5 (2014), 1–16, hier 2.

⁷ BÜCHTER/PALLACK, Methodische Überlegungen, 81.

⁸ Vgl. KÜHN, Steuerung, 109; sowie SVENJA M. KÜHN, Aufgaben in (zentralen) Abschlussprüfungen. Theoretische und empirische Perspektiven auf ein interdisziplinäres Forschungsfeld, in: STEFAN KELLER/CHRISTIAN REINTJES (Hrsg.), Aufgaben als Schlüssel zur Kompetenz. Didaktische Herausforderungen, wissenschaftliche Zugänge und empirische Befunde. Münster 2016, 73–89, hier 77.

Schlagwort *teaching to the test* zudem eine inhaltliche und methodische Einführung von Lerninhalten auf Prüfungsformate.⁹ Während dieses Phänomen in Bezug auf den alltäglichen Unterricht durchaus auch zu einer Verengung des Bildungsbegriffs führen kann, ist eine Eingrenzungen des Stoffes angesichts der Wissensfülle der unterschiedlichen Fachdisziplinen in Hinblick auf Abschlussprüfungen als sinnvoll einzuschätzen.

Da es trotz der Einführung des Zentralabiturs auf Länderebene noch zu Ungleichheiten zwischen den Bundesländern kommt, was beispielsweise ein Längsschnittvergleich der durchschnittlichen Abiturnoten deutlich macht (vgl. Kap. 5.1), ist man bestrebt, die Standardisierung der Abschlussprüfungen noch weiter voranzutreiben. Das föderale Bildungssystem soll – mit dem Argument, ein Wettbewerb zwischen den Bundesländern würde die Qualität der Bildung steigern – durch diese Entwicklungen jedoch nicht grundsätzlich in Frage gestellt werden.¹⁰ Die Ständige Kultusministerkonferenz hat daher im März 2012 den Aufbau eines *Gemeinsamen Abituraufgabenpools der Länder* für die Kernfächer (Mathematik, Deutsch, Englisch und Französisch) beschlossen, welcher auf Grundlage der Bildungsstandards Aufgaben bereitstellen soll, aus denen die Länder seit 2017 eine Auswahl für ihr Abitur treffen können.¹¹ Mit dieser Maßnahme minimiert man in den Kernfächern das Problem der unterschiedlichen Anzahl und Schwierigkeitsgrade der Aufgabensets. Auch die Erwartungshorizonte werden auf diese Weise in den Kernfächern vereinheitlicht, so dass Anforderungen und Qualitätsmerkmale flächendeckend umzusetzen sind. Kritisch zu hinterfragen ist dabei jedoch der Einschnitt in die Freiheit der Bundesländer, die es ihnen gewährt, die Prüfungen an ihren jeweiligen regionalen Kontext anzupassen (vgl. zu diesem Aspekt Kapitel 9.1).

Derzeit werden die Abituraufgabenstellungen im Fach Evangelische Religionslehre auf Grundlage der landesspezifischen Curricula und der EPA »unter Beteiligung von erfahrenen und fachlich profilierten Lehrkräften« entwickelt – so die Formulierung des BAYERISCHEN STAATSMINISTERIUMS FÜR UNTERRICHT UND KULTUS in einer Antwort auf die Anfrage nach den Erstellungsprozessen der Abituraufgabenstellungen in den Nebenfächern. Ebenso gibt es in Niedersachsen Fachkommissionen zur Erstellung der Abituraufgabenstellungen. »Den Kom-

⁹ Vgl. KÜHN, Steuerung, 46.

¹⁰ Vgl. BERND ALTHUSMANN als Niedersächsischer Bildungsminister in einem 2012 verfassten Schreiben an die Schulen. Abrufbar unter: https://www.nibis.de/nli1/allgemein/gosin/zentralabitur/infos2014/Abiturpruefung_2013-2014.pdf (08.02.2020). Niedersachsen, Bayern, Hamburg, Mecklenburg-Vorpommern, Sachsen und Schleswig-Holstein waren Vorreiter in der Entwicklung gemeinsamer Aufgabenpools.

¹¹ Die Aufgabenpools werden in Zusammenarbeit mit Lehrkräften und Fachdidaktikerinnen und Fachdidaktikern vom INSTITUT ZUR QUALITÄTSENTWICKLUNG IM BILDUNGSWESEN (IQB) entwickelt. Vgl. <https://www.iqb.hu-berlin.de/abitur> (12.2.2020).

missionen«, so heißt es in einem Artikel des NIEDERSÄCHSISCHEN KULTUSMINISTERIUMS, »gehören Lehrkräfte an, die in ihrem Fach besonders qualifiziert sind und z.T. weitere hervorgehobene Funktionen, u.a. als Fachberaterin oder Fachberater in der Niedersächsischen Landesschulbehörde, ausüben. Diese Lehrkräfte verfügen alle über aktuelle Unterrichtserfahrungen [...]«.¹² In Thüringen werden die Abituraufgabenstellungen am THÜRINGER INSTITUT FÜR LEHRERFORTBILDUNG, LEHRPLANENTWICKLUNG UND MEDIEN (ThILLM) von an dieses Institut abgeordneten Lehrkräften entwickelt. Nach der Überprüfung der entwickelten Aufgabenstellungen durch die Kultusministerien, werden die Aufgaben zu den Prüfungsterminen an die Schulen übermittelt.

Die EPA geben vor, dass zu den Aufgabenstellungen Erwartungshorizonte erstellt werden müssen. Diese sollen »eine Beschreibung der von den Schülerinnen und Schülern erwarteten Leistungen einschließlich der Angaben von Bewertungskriterien« leisten.¹³ Es wird betont, dass der Erwartungshorizont der Aufgabe keine Musterlösung darstellt, »sondern [...] explizite Kriterien und Gesichtspunkte [benennt], die für die Bewertung einer Prüfungsarbeit notwendig sind«.¹⁴ Es sollen folglich die fachlichen Bewertungskriterien aufgabenspezifisch und explizit ausgewiesen werden. Die Bewertung stellt laut den EPA eine an den Erwartungshorizonten ausgerichtete, kriterienorientierte Entscheidung dar.¹⁵

Korrigiert und bewertet werden die Abiturprüfungen in allen drei untersuchten Bundesländern von einer unterrichtenden Fachlehrerin bzw. einem unterrichtenden Fachlehrer. Die Zweitkorrektur wird in der Regel von einer Fachlehrkraft aus der gleichen Schule erstellt, wobei es in Niedersachsen vorgesehen ist, dass es zur Qualitätssicherung nach drei Abiturdurchgängen einen Lehrkraftaustausch zwischen verschiedenen Schulen in den Zusammensetzungen der Prüfungskommissionen geben kann.¹⁶ Während die Verordnungen in

¹² ANDREAS STEIN, »Zentralabitur« in Niedersachsen. Zur Entstehung von Prüfungsaufgaben mit landesweit einheitlichen Aufgabenstellungen, in: SchulVerwaltung Niedersachsen 1 (2020), 22–24, hier 23.

¹³ Vgl. EPA, 19.

¹⁴ Ebd.

¹⁵ Vgl. ebd.

¹⁶ Vgl. zu den Korrekturbestimmungen: THÜRINGER MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT (Hrsg.), Thüringer Schulordnung. Thüringer Schulordnung für die Grundschule, die Regelschule, die Gemeinschaftsschule, das Gymnasium und die Gesamtschule. Vom 20. Januar 1994. Stand: letzte berücksichtigte Änderung: mehrfach geändert durch Artikel 1 der Verordnung vom 23. Mai 2018 (GVBl. 282). NIEDERSÄCHSISCHES KULTUSMINISTERIUM (Hrsg.), Verordnung über die Abschlüsse in der gymnasialen Oberstufe, im Beruflichen Gymnasium, im Abendgymnasium und im Kolleg (AVO-GOBAG) vom 19. Mai 2005 und geändert durch die Verordnung vom 04. September 2018.

Thüringen und Niedersachsen vorsehen, dass sich der Zweitkorrektor dem Erstkorrektor anschließen oder eine eigene Beurteilung anfertigen kann, ist in Bayern festgelegt, dass zwei gesonderte Korrekturen und Bewertungen angefertigt werden (Bayern Gymnasialschulordnung § 51 Absatz 2). In allen drei Bundesländern wird die Endnote vom Prüfungsvorsitzenden festgesetzt, wenn keine Einigung zwischen Erst- und Zweitkorrektur zustande kommt.

3.2 Vorgaben der Einheitlichen Prüfungsanforderungen für das Fach Evangelische Religionslehre

Die EPA haben bildungsadministrativ für die Sekundarstufe II einen besonderen Weisungscharakter, da sie das einzig bundesweit verbindliche Dokument darstellen, während andere inhaltliche Vorgaben wie beispielsweise die Lehrpläne ausschließlich für ein Bundesland gelten. Die EPA für das Fach Evangelische Religionslehre setzten somit in Hinblick auf die in der Studie untersuchte Urteilsfähigkeit in ethischen und religiösen Fragen den Maßstab für die Anforderung in Abiturprüfungen.

In der Fachpräambel der EPA wird Urteilsfähigkeit inhaltlich gefasst und die Bedeutung der Kompetenz stark hervorgehoben: Schülerinnen und Schüler sollen zur eigenständigen Urteilsfindung befähigt werden, was eine argumentative und sachkundige Teilhabe »am gesellschaftlichen Diskurs über Glauben und Leben« ermöglichen soll.¹⁷ Die Schülerinnen und Schüler sind als Subjekte in diesem Prozess zentral. Urteilsfähigkeit wird nicht nur im Sinne von Argumentationsfähigkeit gedacht, sondern als Weg der Orientierungsfindung in der Welt.¹⁸ Erreicht wird die Befähigung zur eigenständigen Urteilsbildung laut den EPA durch eine dialogische Auseinandersetzung mit anderen Positionen, wobei der Religionsunterricht im Vergleich zu anderen Fächern vor allem darauf angelegt sei, religiöse und dezidiert christliche Traditionen sowie theologische Argumentationsformen in diesen Prozess einzubinden.¹⁹ Neben diesen allgemeinen Aussagen über Urteilsfähigkeit benennen die EPA auch konkrete Anforderungen zur Urteilsbildung, die im Folgenden Betrachtung finden.

BAYERISCHES STAATSMINISTERIUM FÜR UNTERRICHT UND KULTUS (Hrsg.), Schulordnung für die Gymnasien in Bayern (Gymnasialschulordnung – GSO) vom 23. Januar 2007.

¹⁷ Vgl. EPA, 6.

¹⁸ Vgl. ebd.

¹⁹ Vgl. ebd.

3.2.1 Urteilsfähigkeit im Kompetenzmodell der Einheitlichen Prüfungsanforderungen

Das den EPA im Fach Evangelische Religionslehre zugrunde liegende Kompetenzmodell beschreibt religiöse Bildung mit Hilfe von fünf Kompetenzen: Wahrnehmungs- und Darstellungsfähigkeit, Deutungsfähigkeit, Urteilsfähigkeit, Dialogfähigkeit und Gestaltungsfähigkeit.²⁰ Diese Fähigkeiten sollen in Bezug auf religiös bedeutsame Bereiche ausgebildet werden. Die fünf Kompetenzen des den EPA zugrunde liegenden Kompetenzmodells weisen eine große inhaltliche Ähnlichkeit zu den Dimensionen religiöser Bildung auf, die eine Expertengruppe des Comenius-Instituts in ihrem Modell *Grundlegende Kompetenzen religiöser Bildung* (CI-Modell) benennen.²¹ Auch die von der Evangelischen Kirche in Deutschland veröffentlichten Texte *Kompetenzen und Standards für den Evangelischen Religionsunterricht der Sekundarstufe I – Ein Orientierungsrahmen* sowie *Kerncurriculum für das Fach Evangelische Religionslehre in der gymnasialen Oberstufe* legen das durch die EPA vertretene Kompetenzmodell zugrunde.²² Darüber hinaus greifen die meisten Lehrpläne dieses Modell auf.²³ Die Ver-

²⁰ Vgl. a. a. O., 8 f.

²¹ FISCHER/ELSENBAST (Redaktion), *Grundlegende Kompetenzen*, 17. Die Autoren des CI-Kompetenzmodells verweisen bei der Entwicklung der fünf Dimensionen auf ULRICH HEMEL (1988) und CHARLES Y. GLOCK (1969). Die im CI-Modell benannte Dimension »Wahrnehmen/Beschreiben« findet sich in den EPA als »Wahrnehmungs- und Darstellungsfähigkeit« wieder, die Dimension »Verstehen/Deuten« als »Deutungsfähigkeit«, die Dimension »Gestalten/Handeln« als »Gestaltungsfähigkeit«, die Dimension »Kommunizieren/Urteilen« und die Dimension »Teilhaben/Entscheiden« überschneiden sich mit den Kompetenzbereichen »Urteilsfähigkeit« und »Dialogfähigkeit« aus den EPA – die Nähe der Kategorien ist evident. Auch zeitlich liegen beide Dokumente nah beieinander. So wurde die Überarbeitung der EPA ebenso 2006, vier Monate nach dem CI-Modell, veröffentlicht.

²² Vgl. EKD (Hrsg.), *Kompetenzen und Standards* sowie DIES. (Hrsg.), *Kerncurriculum für Evangelische Religionslehre*.

²³ So legen auch die Lehrpläne der Bundesländer Niedersachsen und Thüringen, die in der empirischen Untersuchung im zweiten Teil der Arbeit im Fokus stehen, die fünf Kompetenzen der EPA zugrunde (vgl. THÜRINGER MINISTERIUM FÜR BILDUNG, WISSENSCHAFT UND KULTUR [Hrsg.], *Lehrplan für den Erwerb der allgemeinen Hochschulreife. Evangelische Religionslehre*. 2013, 8–10. NIEDERSÄCHSISCHES KULTUSMINISTERIUM (Hrsg.), *Kerncurriculum für das Gymnasium – gymnasiale Oberstufe, die Gesamtschule – gymnasiale Oberstufe, das Berufliche Gymnasium, das Abendgymnasium, das Kolleg. Evangelische Religion*. 2017, 10.

Der Lehrplan des dritten Bundeslands der Stichprobe – Bayern – ist hingegen stark thematisch gedacht und nimmt keinen Bezug auf das fünfgliedrige Kompetenzmodell (vgl. STAATSMINISTERIUM FÜR SCHULQUALITÄT UND BILDUNGSFORSCHUNG MÜNCHEN [Hrsg.], *Lehrplan. Evangelische Religionslehre*. Jahrgang 11/12. Abrufbar unter: <http://www.isb-gym8-lehr>

ständigkeit auf die benannten fünf im Religionsunterricht zu erwerbenden fachspezifischen Kompetenzen kann folglich vor dem Hintergrund der bildungsadministrativen Vorgaben als Konsens angesehen werden.²⁴

Die Konkretion der Kompetenz Urteilsbildung umfasst im Kompetenzmodell der EPA sechs Stichpunkte, welche die Breite der unter Urteilsfähigkeit zu fassenden Anforderungen verdeutlichen (Hervorhebungen der Operatoren K.M.):

- »deskriptive und normative Aussagen unterscheiden
- Formen theologischer Argumentation vergleichen und bewerten
- Gemeinsamkeiten von Konfessionen und Religionen sowie deren Unterschiede erklären und kriteriengeleitet bewerten
- Modelle ethischer Urteilsbildung kritisch beurteilen und beispielhaft anwenden
- die Menschenwürde theologisch begründen und als Grundwert in aktuellen ethischen Konflikten zur Geltung bringen
- im Kontext der Pluralität einen eigenen Standpunkt zu religiösen und ethischen Fragen einnehmen und argumentativ vertreten.«²⁵

Grau hinterlegt wurden in der obigen Auflistung die Operatoren, um zu verdeutlichen, welche Tätigkeiten im Kompetenzbereich gefordert werden. Definiert werden Anforderungen, jedoch keine Qualitätsmerkmale: Was etwa eine kritische Beurteilung ist und wann die beispielhafte Anwendung ethischer Urteilsbildungsmodelle als qualitativ hochwertig angesehen werden kann, wird nicht gesagt. Somit bleiben die EPA in Bezug auf Bewertungskriterien für Urteilskompetenz – ebenso wie in Bezug auf die anderen vier Kompetenzbereiche –

plan.de/contentserv/3.1.neu/g8.de/index.php?StoryID=26176 [02.09.2020]). Im Fachprofil Evangelische Religionslehre werden Bereiche benannt, in welchen religiöse Kompetenz »eine wichtige Rolle spielen« (z. B. persönliche Lebensgeschichte, religiöse Praxis, Politik...), ohne jedoch Kompetenzbereiche zu benennen (vgl. a. a. O.).

²⁴ Der religionspädagogische Forschungsdiskurs über Modelle religiöser Kompetenz ist gleichsam breiter, wenngleich viele Arbeiten das Modell der EPA geprägt haben (vgl. bspw. RICHARD KABISCH, *Wie lehren wir Religion? Versuch einer Methodik des evangelischen Religionsunterrichts für alle Schulen auf psychologischer Grundlage*. Göttingen ⁵1920, insb. 101; sowie die Arbeiten ULRICH HEMELS, insbesondere: *Ziele religiöser Erziehung. Beiträge zu einer integrativen Theorie*. Frankfurt a. M. 1988; oder auch RUDOLF ENGLERT, *Religionspädagogik in der Schule*, in: FRIEDRICH SCHWEITZER/THOMAS SCHLAG (Hrsg.), *Religionspädagogik im 21. Jahrhundert*. Gütersloh 2004, 79–93).

²⁵ EPA, 8 f.

Auch im folgenden Text werden die Operatoren an einigen Stellen grau hinterlegt, um sie zu akzentuieren. Diese Hervorhebungen stammen alle von der Verfasserin der Studie und sind im Original der Aufgaben nicht enthalten.

unspezifisch.²⁶ Die Qualitätsmerkmale und konkreten Beurteilungskriterien müssen folglich von den Erwartungshorizonten geleistet werden.²⁷ Dass die inhaltliche Ausformulierung von Qualitätsmerkmalen und Kompetenzniveaustufen innerhalb der EPA in allen Fächern ein Manko darstellt, wird auch daran deutlich, dass die 2012 von der Kultusministerkonferenz beschlossenen Bildungsstandards für die Allgemeine Hochschulreife die EPA ablösen.²⁸ Die Entwicklung der Bildungsstandards und das Ablösen der EPA gilt jedoch nur für die Fächer Deutsch, Mathematik, Englisch und Französisch – bundesweite Bildungsstandards für die naturwissenschaftlichen Fächer Biologie, Chemie und Physik für die Sekundarstufe II befinden sich derzeit in der Entwicklung.²⁹ Evangelische Religionslehre ist jedoch noch nicht im Blickfeld dieser Entwicklungen.

3.2.2 Urteilsfähigkeit innerhalb der Anforderungsbereiche

In den EPA werden Anforderungsbereiche (AFB) vorgegeben, welche den Schwierigkeitsgrad einer Aufgabe regulieren sollen. Die AFB sind zum einen allgemein, also fächerübergreifend formuliert und werden zum anderen für jedes Fach nochmals spezifiziert. Die fächerübergreifenden AFB sind definiert durch Reproduktion (AFB I), Anwenden (AFB II) sowie problemlösendes Denken

²⁶ Unter dem Kapitel »3.5 Bewertung von Prüfungsleistungen« werden in den EPA allgemeine Bewertungskriterien formuliert, welche jedoch nicht auf die Schwierigkeit kontroverser Fragestellungen eingehen (siehe Anhang II und vgl. EPA, 19–21).

Zur Beschreibung von Urteilsfähigkeit in der EPA vgl. auch: KATHARINA MUTH, Über die Verortung und Bewertung von Einstellungen und Werthaltungen im Urteilsbildungsprozess. Zur religionspädagogischen Debatte mit einem Blick in die Sozialpsychologie, in: THOMAS HELLER (Hrsg.), Religion und Bildung – Interdisziplinär. Leipzig 2018, 515–526, hier 522 f.

²⁷ Das Fehlen konkreter Beschreibungen von Urteilsbildungskompetenz in bildungsadministrativen Vorgaben monieren auch MARIA MROSCHEN und DIETMAR HÖTTECKE für die naturwissenschaftlichen Fächer Biologie, Physik und Chemie, in denen Urteilskompetenz (unter dem Begriff »Bewertungskompetenz«) seit 2004 in den Bildungsstandards für den Mittleren Schulabschluss verankert ist (vgl. MARIA MROSCHEN/DIETMAR HÖTTECKE, Einstellungen und Vorstellungen von Lehrpersonen zum Kompetenzbereich Bewertung der Nationalen Bildungsstandards, in: Zeitschrift für interpretative Schul- und Unterrichtsforschung 1 (2012), 113–145, hier 114 sowie Kapitel 4.1.4). Das Fehlen von konkret ausformulierten Kriterien für die Diagnose der Kompetenzen in bildungsadministrativen Vorgaben kritisieren auch CORINNA HÖBLE und KERSTIN HEUSINGER VON WALDEGGE in Bezug auf den Biologieunterricht (vgl. DIESS., Bewertungskompetenz diagnostizieren, in: Der mathematische und naturwissenschaftliche Unterricht 63 (2010), 428–434, hier 429).

²⁸ STEIN, »Zentralabitur«, 22.

²⁹ A. a. O., 24.

(AFB III). Dabei wird davon ausgegangen, dass die Komplexität und der Grad der Selbstständigkeit mit der Höhe des AFB zunehmen. Um das Schwierigkeitsniveau zwischen den Wahlaufgaben einer Abiturprüfung und über verschiedene Jahrgänge hinweg weitestgehend zu vereinheitlichen, ist vorgegeben, dass der AFB I und der AFB III in einer Abiturprüfung jeweils mit 30 % vertreten sein sollen, während AFB II mit 40 % den Großteil der Gesamtleistung ausmachen soll.³⁰ Die EPA geben vor, dass sich die Prüfungsaufgabe auf alle drei AFB erstrecken muss, »so dass eine Beurteilung ermöglicht wird, die das gesamte Notenspektrum umfasst«³¹.

Urteilskompetenz als eine eigenständige, begründete Positionierung fällt in den AFB III. Für das Fach Evangelische Religionslehre wird dieser folgendermaßen konkretisiert: »Der Anforderungsbereich III umfasst die *selbstständige systematische Reflexion* und das *Entwickeln von Problemlösungen*, um zu eigenständigen Deutungen, Wertungen, Begründungen, *Urteilen* und Handlungsoptionen sowie zu kreativen Gestaltungs- und Ausdrucksformen zu gelangen«³². Da AFB III neben Urteilen auch die grafisch hervorgehobenen Bereiche Reflexion und Problemlösung beinhaltet, ist eine Aufgabe zur Urteilsbildung im schriftlichen Abitur nicht zwingend. Die mit dem AFB III verbundenen geforderten Leistungen werden in einem Anforderungskatalog genauer beschrieben, der in Bezug auf Urteilsbildung folgende zwei Punkte benennt: Erstens »fachspezifische Positionen, Thesen und Probleme mit dem Ziel einer begründeten und überzeugenden Stellungnahme« erörtern sowie zweitens die »eigene Urteilsbildung unter Beachtung biblischer, theologischer und ethischer Kategorien« reflektieren.³³

Die AFB werden durch Operatoren weiter definiert. Durch diese Festlegung soll jede Aufgabenstellung eindeutig einem AFB zugeordnet werden können. Die folgende Tabelle stellt die Operatoren dar, die ihrer Definition nach eine Urteilsbildung fordern.

³⁰ Vgl. EPA, 10 f., sowie die Einheitlichen Prüfungsanforderungen anderer Fächer.

Erklärend zu den AFB wird in den EPA betont: »Obwohl die Anforderungsbereiche definitorisch unterschieden werden, ergeben sich je nach Aufgabenstellung vielfach Übergänge und Überschneidungen« (EPA, 11).

³¹ A. a. O., 11.

³² Vgl. a. a. O., 12. Hervorhebungen im Original.

³³ Vgl. a. a. O.

<i>Operatoren</i>	<i>Definition</i>
<i>Sich auseinandersetzen mit</i>	ein <i>begründetes eigenes Urteil</i> zu einer Position oder einem dargestellten Sachverhalt entwickeln
<i>Beurteilen</i> <i>Bewerten</i> <i>Stellung nehmen</i> <i>Einen begründeten Standpunkt einnehmen</i>	zu einem Sachverhalt unter Verwendung von Fachwissen und Fachmethoden sich <i>begründet positionieren</i> (Sach- bzw. Werturteil)
<i>Erörtern</i>	die Vielschichtigkeit eines Beurteilungsproblems erkennen und darstellen, dazu Thesen erfassen bzw. aufstellen, Argumente formulieren, nachvollziehbare Zusammenhänge herstellen und dabei eine <i>begründete Schlussfolgerung</i> erarbeiten (dialektische Erörterung)
<i>Prüfen</i> <i>Überprüfen</i>	eine Meinung, Aussage, These, Argumentation nachvollziehen, kritisch befragen und auf der Grundlage erworbener Fachkenntnisse <i>begründet beurteilen</i>
<i>Gestalten</i> <i>Entwerfen</i>	sich textbezogen mit einer Fragestellung kreativ auseinander setzen
<i>Stellung nehmen aus der Sicht von...</i> <i>Eine Erwiderung formulieren aus der Sicht von...</i>	eine unbekannte Position, Argumentation oder Theorie aus der Perspektive einer bekannten Position beleuchten oder in Frage stellen und ein <i>begründetes Urteil</i> abgeben
<i>Konsequenzen aufzeigen</i> <i>Perspektiven entwickeln</i>	Schlussfolgerungen ziehen; Perspektiven, Modelle, Handlungsmöglichkeiten, Konzepte u. a. entfalten

Tab. 1: Für Urteilsfähigkeit relevante Operatoren nach den EPA, 14 f. (Hervorhebungen K.M.).

An den Definitionen der Operatoren wird deutlich, dass Abituraufgabenstellung einer Eigenlogik folgen: Der Prüfling muss wissen, dass ein begründetes Urteil von ihm verlangt wird, wenn der Operator »sich auseinandersetzen mit« genutzt wird, da eine Auseinandersetzung in anderen Kontexten nicht zwangsläufig auch eine Urteilsbildung erfordert. Umso wichtiger ist, dass die verwendeten Operatoren den Prüflingen bekannt sind und eine Passung zwischen den in den Erwartungshorizonten formulierten Anforderungen und den verwendeten Operatoren vorliegt. Anderenfalls werden die Anforderungen für den Prüfling intransparent.

Das Wissen um diese Setzungen und Eigenlogiken ist eine wichtige Verstehensvoraussetzung für die Analyse der Abituraufgabenstellungen. Die Vorgaben müssen jedoch auch ihrerseits diagnostische Forderungen einholen. Zentrale Aspekte der Leistungsdiagnostik sind daher nach einem kurzen Exkurs ein letzter wesentlicher Aspekt des Forschungshintergrundes, auf dem die Studie fußt.

3.2.3 Exkurs: Die Perspektive der EKD – Urteilsfähigkeit zur Orientierungsfindung

Neben den staatlichen Vorgaben ist für den Religionsunterricht auch der Standpunkt der Kirche zentral, da diese den Religionsunterricht in seiner Ausrichtung verfassungsrechtlich mitbestimmt. Aus diesem Grund soll in diesem Exkurs die Stellung von Urteilsfähigkeit in den Denkschriften der Evangelischen Kirche in Deutschland (EKD) herausgearbeitet werden. Zwar haben Denkschriften der EKD einen anderen (ideelleren) Anspruch als bildungsadministrative Vorgaben wie die EPA oder wissenschaftliche Abhandlungen. Dass sie jedoch von den Lehrplänen zitiert werden, macht ihre richtungsweisende Kraft für das Selbstverständnis des Faches deutlich.

Die EKD vertritt ein stärker an christlichen Werten orientiertes Urteilsbildungskonzept. So thematisiert sie in ihren Denkschriften zum Religionsunterricht die Orientierungskraft konfessioneller Gebundenheit im Kontext der Offenheit der Lernprozesse. Die vielbeachtete Denkschrift *Identität und Verständigung* formuliert den Anspruch, dass der Religionsunterricht »identifikatorisches Lernen« ermöglicht und so bei der Identitätsbildung hilft.³⁴ Gleichzeitig fordert die Denkschrift die Öffnung des Religionsunterrichts, um andere Auffassungen zu verstehen und um die Verständigung mit anderen Menschen zu fördern.³⁵ Sie betont, dass auf die Offenheit der Lernprozesse zu achten sei und unterschiedliche Positionen in einer angemessenen Gesprächskultur wahrgenommen, reflektiert und geachtet werden sollen.³⁶ »Andererseits«, so heißt es weiter, »hat die Schule im Ganzen und der Religionsunterricht im Besonderen die Aufgabe, Schülerinnen und Schüler zur verbindlichen Übernahme von Sinnrichtungen für das eigene Leben zu ermutigen und damit auch zu religiöser und konfessorischer Kompetenz zu befähigen.«³⁷

In Bezug auf die ethische Urteilsbildung kritisiert die Denkschrift, dass häufig nur einzelne ethische Probleme in den Blick genommen werden, ohne dass die »ethische Grundlagenproblematik« Beachtung findet, welche der Denkschrift zufolge darin liegt, dass das Nachdenken über moralische Probleme immer mehr zunimmt, während es immer schwerer fällt, Antworten auf die Fragen zu finden.³⁸ Sie fordert den Religionsunterricht heraus, in unserer plu-

³⁴ Vgl. EVANGELISCHE KIRCHE IN DEUTSCHLAND (Hrsg.), *Identität und Verständigung. Standort und Perspektiven des Religionsunterrichts in der Pluralität. Eine Denkschrift der EKD*. Gütersloh 1994, 29. Abrufbar unter https://www.ekd.de/download/identitaet_und_verstaendigung_neu.pdf (07.01.2017).

³⁵ Vgl. ebd.

³⁶ Vgl. bspw. a.a.O., 25.

³⁷ Ebd.

³⁸ Vgl. a.a.O., 15.

ralistischen Gesellschaft einen wesentlichen Beitrag zur gesamtschulischen Aufgabe zu leisten »die junge Generation [...] zu befähigen, Sachverhalte realistisch zu analysieren, sie begründet ethisch zu beurteilen, zu vernünftiger Konsensbildung beizutragen, sowie bereit zu sein, wo notwendig, Lebensgewohnheiten zu ändern (Zukunftsdimension der Bildung).«³⁹

Die jüngste EKD-Denkschrift zum evangelischen Religionsunterricht *Religiöse Orientierung gewinnen – Evangelischer Religionsunterricht als Beitrag zu einer pluralitätsfähigen Schule* setzt einen noch stärkeren Akzent auf die zunehmende religiöse und weltanschauliche Pluralität.⁴⁰ Von dieser zunehmenden Pluralität leitet sie einen erhöhten Orientierungsbedarf seitens der Kinder und Jugendlichen ab.⁴¹ Auf diese Herausforderungen – so die Schlussfolgerung der Denkschrift – muss die Schule und speziell der Religionsunterricht reagieren, indem er die Pluralität in reflektierter Form wahrnimmt, sie bearbeitet und so die religiöse Orientierungsfähigkeit der Schüler fördert.⁴² Gemäß dem Titel sollen die Lernenden durch den Religionsunterricht »religiöse Orientierung gewinnen«. Dies soll er leisten, indem er die ethische und religiöse Urteilsbildung fördert: Der Religionsunterricht hilft – so der in der Denkschrift formulierte Anspruch – »die Frage nach dem Ganzen und nach dem tragenden Sinn von allem [...] aufgrund der Förderung des freien und selbstständigen eigenen ethischen und religiösen Urteils zu beantworten.«⁴³ Die freie und selbstständige Urteilsbildung wird als Schlüssel wahrgenommen, Antworten auf existentielle Fragen zu finden.

In Konflikten, welche aufgrund unterschiedlicher religiöser Wertevorstellungen entstehen, hält die Denkschrift eine »kompetente und kritische, zugleich aber verständigungsbereite Bearbeitung religiöser Wertvorstellungen [...] aus pädagogischen Gründen [für] unausweichlich.«⁴⁴ Es wird jedoch nicht genauer spezifiziert, was diese »kompetente Bearbeitung« ausmacht.

Festzuhalten gilt, dass die Denkschriften der EKD Werte- und Urteilsbildung als wesentliche Bestandteile des Religionsunterrichts hervorheben und die Selbstständigkeit der Urteilsbildung betonen. Gleichzeitig werden aber auch Werte definiert, welche der Religionsunterricht vertreten und vermitteln soll. Die Stoßrichtung der Urteilsbildung, wird durch das Erziehungsziel »Pluralitätsfähigkeit« und Wertebildung (Empathie, Toleranz, Respekt etc.) festgesetzt. Die Forderungen der Denkschriften gehen damit über eine rein theoriegeleitete

³⁹ EKD (Hrsg.), Identität und Verständigung, 16.

⁴⁰ Vgl. EKD (Hrsg.), Religiöse Orientierung gewinnen. Evangelischer Religionsunterricht als Beitrag zu einer pluralitätsfähigen Schule. Gütersloh 2014.

⁴¹ Vgl. a. a. O., 11.

⁴² Vgl. ebd.

⁴³ A. a. O., 8 f.

⁴⁴ A. a. O., 22.

Reflexion hinaus, indem sie nicht das Anwenden ethischer Theorien, sondern Wertebildung und Orientierung ins Zentrum rücken.

Dass die EKD einer Orientierung-gebenden Urteilskompetenz im Religionsunterricht einen hohen Stellenwert beimisst, wird auch in den von ihr formulierten Kompetenzen für die Religionslehrerausbildung deutlich. In diesen wird hervorgehoben, dass es das Ziel der Gestaltung von Lehr- und Lernprozessen sein soll, Schülern und Schülerinnen »den Erwerb von Kompetenzen zu ermöglichen, die ihnen helfen, sich in der religiös pluralen Welt zu orientieren, eigene religiöse Überzeugungen zu gewinnen, darüber auskunfts- und dialogfähig zu sein sowie ethisch verantwortlich urteilen und handeln zu können.«⁴⁵ In den Standards wird wiederholt betont, dass der Unterricht »aus der Perspektive des christlichen Glaubens und im Horizont evangelischer Freiheit gestaltet« werden soll.⁴⁶ In einem weiteren Standard wird unter der Überschrift »Religionslehrerinnen und Religionslehrer nehmen die Erziehungsaufgabe vor dem Hintergrund eines theologisch reflektierten, christlichen Menschenbildes bewusst wahr« das Ziel formuliert, »die Gewissen der Schülerinnen und Schüler zu schärfen, ethische Orientierungen aus christlicher Perspektive zu vermitteln und die Fähigkeit und Bereitschaft zu wecken, Glaubensüberzeugungen zu achten, Toleranz zu üben und Solidarität zu praktizieren«.⁴⁷

Die EKD macht folglich deutlich, dass es ihr um die Förderung eigenständiger Urteilsbildung geht. Es geht ihr nicht um eine Übernahme des Bekenntnisses, auf dem der Religionsunterricht fußt, sondern um eine Auseinandersetzung mit diesem Bekenntnis.

3.3 Diagnostische Grundlagen und Besonderheiten zur Urteilsfähigkeit im Kontext von Abschlussprüfungen

3.3.1 Formen von Kompetenzmodellen

Zunächst ist zu klären, welche Voraussetzungen für die Diagnostik bestehen: Diagnostik verfolgt den Zweck, »individuelle Verhaltensmuster nach einem objektiven Schema zu erheben und zu bewerten«.⁴⁸ Dies setzt ein konkretes Konstrukt von dem zu prüfenden Verhaltensmuster voraus. Die Kompetenzorien-

⁴⁵ EKD (Hrsg.), Theologisch-Religionspädagogische Kompetenz. Professionelle Kompetenzen und Standards für die Religionslehrerausbildung. Empfehlungen der Gemischten Kommission zur Reform des Theologiestudiums. Hannover 2008, 18.

⁴⁶ Vgl. ebd.

⁴⁷ Vgl. a. a. O., 19.

⁴⁸ Vgl. REIS/SCHWARZKOPF, Diagnose religiöser Lernprozesse, 42.

tierung fragt dezidiert nach dem Aufbau der Fähigkeit, die geprüft werden soll. NICLAS SCHAPER unterscheidet drei Formen von Kompetenzmodellen: Kompetenzstrukturmodelle, Kompetenzniveaumodelle und Kompetenzentwicklungsmodelle.⁴⁹ Kompetenzstrukturmodelle fragen nach den Facetten »personeller Ressourcen (Kenntnisse, Fertigkeiten, Fähigkeiten etc.)«, welche zur »Bewältigung situationaler Anforderungen einer Domäne erforderlich sind und wie diese Facetten zusammenhängen«.⁵⁰ Dies ermöglicht, die Binnenstruktur einer Kompetenz differenziert beschreiben zu können. »Hierbei geht man davon aus, dass die Teilkompetenzen zwar einerseits miteinander korreliert sind. Andererseits sollten aber für eine differenzierte Kompetenzbeschreibung und -messung verschiedene – hinreichend unabhängige – Dimensionen herangezogen werden.«⁵¹ Sogenannte Kriterienraster arbeiten mit den Teilkompetenzen und legen damit implizit Kompetenzstrukturmodelle zugrunde.

Kompetenzniveaumodelle haben eine inhaltliche Ausrichtung und differenzieren verschiedene Ausprägungsgrade einer Kompetenz. Diese Modelle ermöglichen eine Einordnung in ein nach Stufen hierarchisch geordnetes Niveau. Es geht dabei nicht um eine Einordnung der Leistung in Pole wie »entspricht der Leistung« oder »entspricht der Leistung nicht«, sondern um »Kompetenzniveaucharakterisierungen«, welche erwartbare Handlungen objektiv beschreiben.⁵²

Die dritte Gruppe, die Kompetenzentwicklungsmodelle, benennen eine Abfolge der Kompetenzaneignung. Sie treffen folglich Aussagen über den Kompetenzerwerbsprozess und darüber, welche Ausprägung einer Kompetenz »auf den jeweiligen Lern- bzw. Entwicklungsstufen vermittelt werden sollten bzw. erwartbar sind«.⁵³

An einem Beispiel für Urteilskompetenz aus der Biologiedidaktik von ELKE VISSER, welches im folgenden Kapitel noch expliziter dargestellt wird, kann gut illustriert werden, dass alle drei Modellformen ineinandergreifen: ELKE VISSER modelliert Urteilskompetenz in drei Bereichen: Sie differenziert in deskriptive Aspekte (Perspektiven- und Folgeberücksichtigung), normative Aspekte (Normen und Werte der Begründung) und argumentative Aspekte (vgl. Kapitel

⁴⁹ Vgl. SCHAPER, Kompetenzmodellierung, 173; sowie RÜDIGER PREIBER, Kompetenzorientierte Hochschuldidaktik, in: FLORIAN BRUCKMANN/OLIVER REIS/MONIKA SCHEIDLER (Hrsg.), Kompetenzorientierte Lehre in der Theologie. Konkretion – Reflexion – Perspektiven. Berlin 2011, 17–36, hier 22.

⁵⁰ Vgl. SCHAPER: Kompetenzmodellierung, 174.

⁵¹ Ebd.

⁵² Vgl. a. a. O., 175.

⁵³ A. a. O., 176.

4.1.4).⁵⁴ Mit dieser Unterteilung ist die Kompetenzstruktur benannt. Jeder dieser drei Aspekte bedarf wiederum einer Kompetenzniveaucharakterisierung, damit messbar wird, auf welchem Niveau man den Aspekt beherrscht. So bestimmt VISSER für deskriptive Aspekte, drei Niveaustufen: Auf Niveaustufe I werden keine Folgen oder nur Folgen für unmittelbar Beteiligte benannt. Auf Stufe II werden Folgen auch für andere als die der unmittelbar Betroffenen benannt und auf Niveaustufe III wird die Wirkung auf gesellschaftlicher Ebene bei der Folgenreflexion mitberücksichtigt.⁵⁵ Aus diesem Modell ließe sich nun ein Kompetenzentwicklungsmodell ableiten, welches Lern- und Entwicklungsstufen zur jeweiligen Altersgruppe modelliert.

3.3.2 Bezugsnormen, Gütekriterien und Bewertungsansätze

Durch Kompetenzmodelle wird die zu prüfende Kompetenz anhand fachlicher Kriterien zu beschreiben versucht, also eine kriteriale oder auch *sachliche Bezugsnorm* zugrunde gelegt, um die Vergleichbarkeit zwischen den Bewertungen und damit letztendlich zwischen den Abschlüssen zu gewähren.⁵⁶ Die Vergleichbarkeit ist an die Testgütekriterien Objektivität, Reliabilität und Validität geknüpft. Objektivität bestimmt dabei den Grad der Unabhängigkeit der Durchführung, Auswertung und Interpretation eines Tests von der durchführenden Person.⁵⁷ Mit Blick auf die Bewertung von Urteilsfähigkeit in schriftlichen Abiturprüfungen kommt dabei vor allem die Auswertungsobjektivität in den Blick. Diese ist gegeben, wenn die Bewertung von der bewertenden Person unabhängig ist – folglich, wenn mehrere Korrektorinnen und Korrektoren zu der

⁵⁴ Vgl. ELKE VISSER, Die Diagnose der Bewertungskompetenz durch schriftliche Aufgaben im Biologieunterricht. Hamburg 2014.

⁵⁵ Vgl. a. a. O., 285–295.

⁵⁶ Man unterscheidet neben der kriterialen Bezugsnorm noch die individuelle Bezugsnorm, welche den individuellen Lernzuwachs der Schülerinnen und Schüler berücksichtigt sowie die nicht zulässige, da über Klassengrenzen hinweg nicht aussagekräftige, soziale Bezugsnorm, welche den Leistungsstand der besten einer Lerngruppe zur Bezugsnorm erhebt. Vgl. zum Feld der Notengebung im Religionsunterricht: THORSTEN BOHL, Schulische Notengebung. Probleme und Entwicklungsmöglichkeiten, in: Katechetische Blätter 132 (2007), 249–254.

⁵⁷ Vgl. MATTHIAS SCHWAIGHOFER/MORITZ HEENE/MARKUS BÜHNER, Grundlagen und Kriterien der Diagnostik, in: DETLEF URHAHNE/MARKUS DRESSEL/FRANK FISCHER (Hrsg.), Psychologie für den Lehrberuf. Berlin/Heidelberg: 2019, 493–516, hier 476; sowie ebenso zu den Testgütekriterien WERNER SACHER, Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe. Bad Heilbrunn ⁵2009, 36–42.

gleichen Beurteilung der Leistung kommen.⁵⁸ »Für eine hohe Objektivität sind genaue Auswertungsregeln notwendig.«⁵⁹

Das Kriterium Validität beschreibt, ob ein Test tatsächlich das misst, was er zu messen beansprucht.⁶⁰ In Bezug auf Urteilsfähigkeitsaufgaben im schriftlichen Abitur kann beispielsweise die Textlänge des der Aufgabenstellung zugrunde liegenden Materials angesichts der vorgegebenen Bearbeitungszeit zu einem ergebnisbedingenden Faktor werden. So nimmt die Validität ab, wenn beispielsweise Lesekompetenz das Testergebnis beeinflusst, obwohl über diese Kompetenz keine Aussage getroffen werden soll.

Reliabilität beschreibt die Messgenauigkeit, also die Zuverlässigkeit, mit der ein Test eine Fähigkeit misst.⁶¹ Reliabilität kann unabhängig von der Validität gegeben sein. So kann ein Test wiederholt zu ähnlichen Ergebnissen führen, aber dennoch andere Kompetenzen prüfen, als die von ihm intendierten. Der Test wäre dann reliabel, aber nicht valide.⁶²



Abb. 4: Die klassischen Gütekriterien der Testtheorie.

Die drei benannten Gütekriterien gelten als die *klassischen Gütekriterien der Testtheorie* und sind vor allem im wissenschaftlichen Kontext bindend. In alltäglichen Diagnose-, Prüfungs- und Bewertungsverfahren in der Schule sind sie als anzustrebende Ziele zu betrachten. Mit Blick auf zentrale Abschlussprüfungen haben sie wiederum eine hervorgehobene Rolle, da Abschlussprüfungen zur Selektion und Allokation von zukünftigen Berufslaufbahnen und Positionsverteilungen wesentlich beitragen.⁶³

⁵⁸ Vgl. SCHWAIGHOFER/HEENE/BÜHNER, Grundlagen Diagnostik, 476.

⁵⁹ Ebd.

⁶⁰ Vgl. a. a. O., 478.

⁶¹ Vgl. a. a. O., 477.

⁶² Vgl. ebd. Reliabilität ist jedoch eine Voraussetzung für Validität: Ist die Reliabilität gering, so kommt es zu fehlerbehafteten Messwerten, was auf eine geringe Validität der Indikatoren für die Messung deuten kann (vgl. a. a. O., 480).

⁶³ Vgl. zu den von HELMUT FEND herausgearbeiteten gesellschaftlichen Funktionen von Schule (Qualifikationsfunktion, Allokationsfunktion, Integrationsfunktion, Enkulturations-

Abschlussprüfungen sind dem Bereich der Lernstandsdiagnostik zuzuordnen, welcher auch als summative Leistungsmessung von einer formativen Leistungsmessung abgegrenzt wird, die im Sinne einer Lernprozessdiagnostik über den reinen Stand des Lernens hinaus nach Möglichkeiten der weiteren Förderung und nächsten Entwicklungsstufen fragt. »Formative assessment«, so definieren DANIELLE COLARDYN und JENS BJORNAVOLD die weitverbreitete Differenzierung von Leistungsbewertungen, »is input-driven, centred on the education and training procedure and linked to educational standards. Summative assessment is outcome-driven, centred on results achieved and linked to occupational standards (non-formal and informal learning) or to educational standards (formal learning).«⁶⁴ Auf dieser Differenzierung basierend unterscheiden REIS und SCHWARZKOPF die Kriterien der Messung in vertikale Dimensionen, welche in Bezug auf die summative Lernstandsdiagnostik von Relevanz sind und horizontale Dimensionen, welche sich auf die formative Lernprozessdiagnostik beziehen⁶⁵: Während Lernprozessdiagnostik verschiedene Tätigkeitsmerkmale im Lernprozess als Kriterien betrachtet, soll Lernstandsdiagnostik verschiedene Qualitäten der Bewältigung einer bestimmten Aufgabenbearbeitung zu einem bestimmten Zeitpunkt erfassen und benötigt daher eine nach Ausprägungsgraden hierarchisch geordnete Stufung der Qualität als Maßstab.⁶⁶

funktion): DERS., *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen*. Wiesbaden ²2008, 51.

⁶⁴ DANIELLE COLARDYN/JENS BJORNAVOLD, *Validation of Formal, Non-Formal and Informal Learning. Policy and practices in EU Member States*, in: *European Journal of Education* 39 (2004), 69–89, hier 79 f.

⁶⁵ Vgl. REIS/SCHWARZKOPF, *Diagnose religiöser Lernprozesse*, 77.

⁶⁶ Vgl. a. a. O., 77. Lernprozessdiagnostik intendiert nach REIS und SCHWARZKOPF, den Lernprozess sichtbar zu machen, indem er hinsichtlich der Normenbearbeitung gelesen wird. REIS und SCHWARZKOPF schlagen dafür ein Modell vor, welches die Tätigkeitsmerkmale im Prozess als Kriterien fasst und prüft, für welche Optionen sich die Lernenden entscheiden (vgl. REIS/SCHWARZKOPF, *Diagnose religiöser Lernprozesse*, 77). Anstatt die »vertikale Organisation von Ausprägungsmerkmalen des Learning-Outcomes« zu differenzieren, geht es bei der Lernprozessdiagnostik um »eine horizontale Rekonstruktion der Bearbeitung entlang von Entscheidungsoptionen«, welche durch eine Bearbeitungsstruktur möglich wird (vgl. a. a. O., 78).

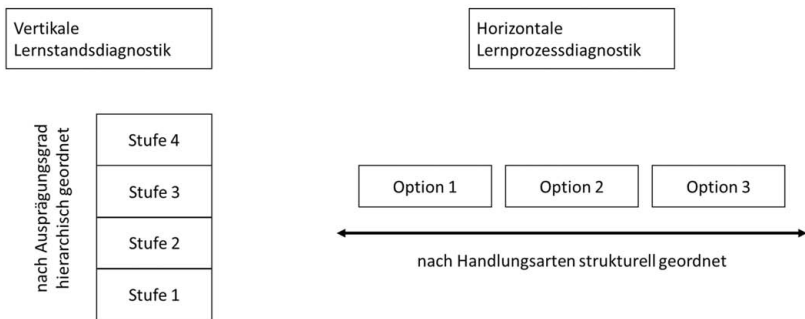


Abb. 5: Vertikale Lernstandsdiagnostik und horizontale Lernprozessdiagnostik. Nachbildung aus REIS/SCHWARZKOPF: Diagnose religiöser Lernprozesse, 77.

Um die notwendige Hierarchisierung auszudifferenzieren, haben BERND SCHÖNEMANN, HOLGER THÜNEMANN und MEIK ZÜLSDORF-KERSTING im Kontext einer Studie zur Evaluation von Abituraufgabenstellungen im Fach Geschichte eine Systematik zur Leistungsbeschreibung entwickelt, in der sie drei Modi bestimmen⁶⁷: Im *binären Modus* wird ein Kriterium nach den Polen »erbracht«/»nicht-erbracht« bewertet, während ein *gradueller Modus* unterschiedliche Stufen der erbrachten Leistung unterscheidet. Ein dritter Modus, der von den Autoren als *systematischer Modus* benannt wird, erfasst Leistungen, die in unterschiedlichen Bereichen erbracht werden, beispielsweise, wenn »ein Problem inhaltlich, theoretisch, terminologisch oder methodologisch reflektiert« wurde.⁶⁸

Eine binäre Differenzierung ist in erster Linie im AFB I sinnvoll, wenn man prüft, ob etwas genannt oder nicht genannt wurde. In Bezug auf komplexe Leistungen sind binäre Differenzierung nur hinsichtlich übergeordneter Qualitätsmerkmale sinnvoll, die jedoch noch weiter ausdifferenziert werden müssen: Wenn beispielsweise ein Erwartungshorizont für eine Urteilsbildungsaufgabe eine Stellungnahme oder eine Begründung fordert, kann binär entschieden werden, ob diese vorhanden oder nicht vorhanden ist. Das Vorhandensein ist für die Lösung jedoch nur eine notwendige, aber keine hinreichende Bedingung. Die Begründung muss weiteren Anforderungen entsprechen, um etwas über die Qualität der Lösung aussagen zu können. Die Frage, was die Qualität der Lösung bestimmt, ist abhängig vom wissenschaftlichen Diskurs. So kann nach GEORG LIND eine Begründung, die unterschiedliche Positionen mitbedenkt und kritisch reflektiert als qualitativ hochwertiger angesehen werden als eine lineare, aus-

⁶⁷ Vgl. BERND SCHÖNEMANN/HOLGER THÜNEMANN/MEIK ZÜLSDORF-KERSTING, Was können Abiturienten? Zugleich ein Beitrag zur Debatte über Kompetenzen und Standards im Fach Geschichte. Berlin/Münster 2011, 34.

⁶⁸ Vgl. ebd.

schließlich konformistische Begründung (vgl. Kapitel 4.1.2). Zur Erfassung dieser Qualitäten braucht es eine graduelle Beurteilung, welcher die Definition verschiedener Niveaustufen zugrunde liegen muss – die sogenannten Niveaustufenmodelle (vgl. Kapitel 3.3.1). Um bewertbar zu sein, müssen die Niveauabstufungen aller Teilkompetenzen »durch Deskriptoren beschrieben werden, mit deren Hilfe die Anforderungen auf den unterschiedlichen Niveaus in praktisch beobachtbare oder abfragbare Handlungen transferiert werden.«⁶⁹ (Beispiele für ein Niveaustufenmodell zur Urteilsfähigkeit siehe Kapitel 4).

Die Bewertung anhand von Niveaustufenmodellen wird auch von der Klieme-Expertise vorgesehen. Dort heißt es:

»Jede Kompetenzstufe ist durch kognitive Prozesse und Handlungen von bestimmter Qualität spezifiziert, die Schülerinnen und Schüler auf dieser Stufe bewältigen können, nicht aber auf niedrigeren Stufen.«⁷⁰

Um eine gute Bewertungsgrundlage bieten zu könnten, müssten die Erwartungshorizonte von Abituraufgabenstellungen aus kompetenzorientierter Sicht eine Verständigung über graduelle Abstufungen eines Qualitätsmerkmals ermöglichen.

Eine letzte für diese Studie zentrale Unterscheidung diagnostischen Vorgehens liegt in dem Unterschied zwischen *synthetischer* und *analytischer Bewertung*. Beide Formen haben Stärken und Grenzen: Eine synthetische Bewertung geht von einem Gesamteindruck aus, welcher im Anschluss unter Berücksichtigung verschiedener Perspektiven auf Details differenziert und überprüft wird.⁷¹ Eine analytische Bewertung richtet den Fokus auf einzelne Kategorien, welche bewertet und dann zu einem Gesamturteil zusammengesetzt werden.⁷² Analytische Beurteilungen haben im Zuge der Kompetenzorientierung eine starke Aufwertung erfahren, da sie eine höhere Objektivität versprechen. Sie werden oft in Form von Kriterienrastern umgesetzt. Die Stärke dieser Raster ist, dass die Teilaspekte einer Leistung systematisiert und durch vorgegebene Punkte gewichtet werden. Dies beugt Messfehlern vor und steigert die Validität: So kann beispielsweise eine Schwäche in einem Teilbereich nur durch den Abzug der für diesen Teilbereich vorgesehenen Punkte in die Gesamtbewertung einfließen, während beispielsweise ein schlechter Ausdruck bei einer synthetischen Beurteilung schnell den Gesamteindruck negativ dominieren kann. Die Herausforderung analytischer Bewertungsansätze besteht darin, die Entität des Ganzen im Blick zu behalten. Ein gewichtiger Einwand gegen analytische Beurteilungen

⁶⁹ PREIßER, Kompetenzorientierte Hochschuldidaktik, 22.

⁷⁰ Vgl. BMBF (Hrsg.), Bildungsstandards, 76.

⁷¹ Vgl. SACHER, Leistungen, 151.

⁷² Vgl. a. a. O., 152.

bezieht sich auf Fehlgewichtungen, durch welche eine gute Note (beispielsweise durch die Einhaltung formaler Grundsätze) erreicht werden kann, ohne den Inhalt tiefgründig bearbeitet zu haben. Zum anderen können komplexe Aufgabenstellungen so verschiedenartige Lösungen evozieren, dass Teile dieser durch die vorgegebenen Raster nicht erfasst werden. Synthetische Bewertungen haben diese Aspekte eher im Blick. Sie laufen jedoch Gefahr, für die Lösung zentrale Aspekte nicht zu berücksichtigen oder einigen Aspekte unbewusst einen höheren Stellenwert einzuräumen als es angemessen wäre. Zudem können die Kenntnisse und Fähigkeiten der Lehrkraft, die domänenspezifisch sind, in diesem Ansatz schneller zum Maßstab der Bewertung werden: Der »Kontrast-/Ähnlichkeitsfehler« als ein Messfehler, bei dem die Lehrkraft von sich selbst auf die Schülerinnen und Schüler schließt, wird wahrscheinlicher.⁷³

WERNER SACHER spricht in Bezug auf »ganzheitliche Leistungen«, unter denen er nicht leicht operationalisierbare Leistungen wie beispielsweise einen Aufsatz oder eine bildnerische Leistung fasst, von einer *übersummativen Qualität*, »die sich nicht aus Punkten und der Erfüllung einzeln angelegter Kriterien hochrechnen lässt.«⁷⁴ Er schlägt für die Bewertung ganzheitlicher Leistungen einen hermeneutischen Kreisprozess vor, bei welchem von einem anfänglich undifferenzierten Gesamteindruck ausgegangen wird, welcher dann anhand von vorab festgelegten Kriterien mit ihren entsprechenden Ausprägungsgraden überprüft und korrigiert wird.⁷⁵ Zum Schluss sollen die überprüften Kriterien wieder in einen »abschließend-differenzierten Gesamteindruck« münden.⁷⁶ SACHER spricht sich dafür aus, die Liste der Bewertungskriterien als unabgeschlossen aufzufassen und »ein oder zwei Kriterien zuzulassen, die angesichts der konkreten Schülerleistung ad hoc formuliert und angewendet werden können«.⁷⁷ Dies wahrt die Offenheit der Beurteilung und gibt den korrigierenden Lehrkräften Spielraum, auf die Besonderheiten einer komplexen Leistung einzugehen.

⁷³ Vgl. zum Kontrast-/Ähnlichkeitsfehler MATTHIAS VON SALTERN, Schulleistung 2.0. Von der Note zum Kompetenzraster. Norderstedt 2011, 101.

⁷⁴ SACHER, Leistungen, 155.

⁷⁵ Vgl. ebd.

⁷⁶ Vgl. ebd.

⁷⁷ A.a.O., 153.

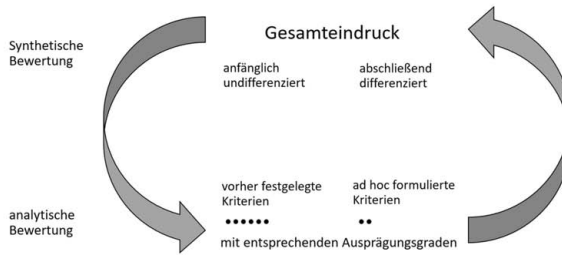


Abb. 6: Hermeneutischer Kreisprozess zur Beurteilung ganzheitlicher Leistungen. Eigene Darstellung nach SACHER: Leistungen entwickeln, überprüfen und bewerten, 155.

3.3.3 Urteilsbildung im Medium der Schriftlichkeit

Neben den im Vorangegangenen beschriebenen allgemeinen Aspekten der Leistungsbeurteilung sind mit der im Fokus stehenden Urteilskompetenz Besonderheiten hinsichtlich ihrer Prüfung verbunden: So ist das Medium der Schriftlichkeit, in dem die untersuchten Abiturprüfungen Urteilsbildungen verlangen, für die Aufgabenlösung nicht unerheblich. Interaktionsdimensionen von Diskursen können im schriftlichen Abitur nicht geprüft werden, so dass Kompetenzen wie das spontane Reagieren auf Gegenargumente und das Pflegen einer wertschätzenden Streitkultur – die im Medium des Mündlichen ein Teil des Urteilsbildungsprozesses sind – nicht überprüft werden können. Beim Schreiben sind im Vergleich zur Mündlichkeit eine höhere Sorgfalt und Reflexivität zu erwarten.⁷⁸ Die Strukturierung und Begründung von Argumenten rücken in den Vordergrund schriftlicher Urteilsbildungsaufgaben. Die Schriftlichkeit verschiebt dabei »die Argumentation als Prozess hin zur Argumentation als Produkt«, was zur Folge hat, dass die originär dialogische Form hinter die monologische Form von Schreibprozessen rückt.⁷⁹ Zwar kann eine schriftliche Argumentation als Diskursbeitrag aufgefasst werden, jedoch müssen etwaige Gegenargumente von der Verfasserin bzw. vom Verfasser selbst vorgebracht und entkräftet werden, was einen Perspektivwechsel notwendig macht.⁸⁰ Neben der Produktion von Diskursbeiträgen, welche vor allem bei Gestaltungsaufgaben zum Tragen kommt, gibt es »Textaufgaben« und »erweiterte Textaufgaben«, die einen Akzent auf die Rezeption von Argumenten legen (vgl. zur Beschreibung der

⁷⁸ Vgl. ELKE VISSER/CORINNA HÖBLE, Bioethisch argumentieren – Ein Diagnostischer Blick auf die Bewertungskompetenz im Biologieunterricht, in: ALEXANDRA BUDKE/MIRIAM KUCKUCK/MICHAEL MAYER/FRANK SCHÄBITZ/KIRSTEN SCHLÜTER/GÜNTHER WEISS (Hrsg.), Fachlich argumentieren lernen. Didaktische Forschungen zur Argumentation in den Unterrichtsfächern. Münster/New York 2015, 182–195, hier 183.

⁷⁹ Ebd.

⁸⁰ Vgl. ebd.

Aufgabenarten auch Kap. 6.2.1).⁸¹ Bei diesen Aufgaben steht die Rekonstruktion von Argumenten und die kritische Prüfung dieser im Zentrum.

Nachdem in diesem Kapitel die zentralen Kontexte beschrieben und diskutiert wurden, in welche Leistungsbewertungen in zentralen Abituraufgabenstellungen eingebettet sind, widmet sich das folgende Kapitel der Fragestellung, welche Qualitätsmerkmale zur Messung von Urteilskompetenz im aktuellen Forschungsdiskurs diskutiert werden.

⁸¹ Vgl. zu den Aufgabenarten EPA, 16–18.

4 Qualitätsmerkmale ethischer und religiöser Urteilskompetenz – Aktueller Forschungsstand

Das folgende Kapitel zielt darauf ab, Vorschläge aus der Forschungsliteratur zur Messung von Urteilsqualität herauszuarbeiten. Die Darstellung des aktuellen Forschungsstandes über Qualitätsmerkmale ethischer und religiöser Urteilskompetenz gliedert sich dabei in zwei Teile: Im ersten Teil werden allgemeinpädagogische Forschungen sowie Studien anderer – nicht religionspädagogischer – Fachdidaktiken vorgestellt, während der zweite Teil religionspädagogische Forschungen betrachtet. Wie bereits im vorangegangenen Kapitel deutlich wurde, ist Urteilsbildung im Medium des Schriftlichen eng mit Argumentationsfähigkeit verknüpft. Aus diesem Grund steigt die Darstellung des Forschungsstandes mit einem Blick in die Argumentationstheorie ein, welche formale Aspekte des Argumentierens und Urteilens untersucht und beschreibt (Kap. 4.1.1). Darauf folgt eine inhaltliche, entwicklungspsychologische Perspektive auf Urteilsfähigkeit in Kapitel 4.1.2, welches sich mit Forschungen von LAWRENCE KOHLBERG und GEORG LIND befasst und danach fragt, welche Qualitätsmerkmale aus diesen Forschungen zu ziehen sind. In Kapitel 4.1.3 und Kapitel 4.1.4 werden empirische Forschungen vorgestellt, welche formale Aspekte der Argumentationstheorie mit inhaltlichen Aspekten von Urteilsfähigkeit vereinen und Kompetenzniveaustufen benennen.

Die Darstellung des religionspädagogischen Forschungsstandes zum Bereich Urteilsbildung im zweiten Teil des Kapitels nimmt seinen Ausgangspunkt in rechtlichen Fragen nach der Stellung des konfessionell gebundenen Religionsunterrichts in unserem zur religiös-weltanschaulichen Neutralität verpflichteten Staat. Dieser rechtliche Zugang mag etwas ungewöhnlich für die Darstellung eines Forschungsstandes erscheinen, jedoch ist die Frage nach der Bekenntnisorientierung im Kontext von Urteilsfähigkeit für das Verständnis des Forschungsdiskurses und für das Verständnis der Spezifika von Qualitätsmerkmalen religiöser Urteilsbildung zentral. Aus ihnen leiten sich verschiedene Ansichten über den Umgang mit der konfessionellen Bestimmtheit in Urteilsbildungsprozessen ab, welche in Kapitel 4.2.1 vorgestellt werden.

Eine lange Tradition haben in Theologie und Religionspädagogik Modelle ethischer Urteilsbildung, welche das Urteilen als einen Prozess mit unter-

schiedlichen Bearbeitungsschritten fassen. Die in diesen Bearbeitungsschritten implizit enthaltenen Qualitätsansprüche an Urteilsbildungen werden in Kapitel 4.2.2 besprochen. Im Anschluss daran beleuchtet Kapitel 4.2.3 religionspädagogische empirische Studien zur Argumentations- und Urteilsfähigkeit von Schülerinnen und Schülern ebenso vor dem Hintergrund der Fragestellung nach Qualitätsmerkmalen. Abschließend werden die im Forschungsdiskurs als zentral erachteten Qualitätsmerkmale von Urteilsbildung zusammengefasst (Kapitel 4.3).

4.1 Allgemeinpädagogische Forschungen

4.1.1 Argumentationstheorie

Die Qualität von Urteilsbildungen bestimmt sich wesentlich durch die Begründung, welche – wie im vorangegangenen Kapitel 3.3.3 gezeigt wurde – im Medium des Schriftlichen noch stärker an formalen Aspekten ausgerichtet ist als im mündlichen Diskurs. Mit dem Austausch von Gründen bezüglich einer strittig gewordenen Aussage beschäftigt sich die Argumentationstheorie.¹ Ziel der Argumentation ist, die Strittigkeit zu bearbeiten und Geltung von Argumenten herzustellen.² Argumentieren zeichnet sich durch Verhandelbarkeit aus, im Gegensatz zum Erklären, für welches Gewissheit konstitutiv ist. Im Kontext der Urteilsfähigkeit dient die Argumentation der Begründung des über eine kontroverse Fragestellung gefällten Urteils.

Ein Argument – darin stimmen die vielfältigen Argumentationstheorien überein – besteht aus einer Behauptung oder einer Aussage, welche durch Rechtfertigungen, Begründungen oder Stützungen untermauert wird.³ Seit den Anfängen der Rhetorik wird das Argument als ein dreiteiliges Modell gedacht, dessen drei Aussagen unterschiedliche Funktionen bedienen: Es gibt einen »Grund«, der die Position bezeichnet, von der aus man etwas begründet und eine »Konklusion«, die begründet wird. Der »Übergang« koppelt dabei Grund und Konklusion.⁴ Je nach Theorie wird der Übergang als Schlussregel, Oberprämisse,

¹ Vgl. KATI HANNKEN-ILLJES, *Argumentation. Einführung in die Theorie und Analyse der Argumentation*. Tübingen 2018, 19.

² Vgl. a. a. O., 21.

³ Vgl. THOMAS WEIß/NICOLAI BASEL/UTE HARMS/HELMUT PRECHTL/MARTIN ROTHGANGEL, *Fächerübergreifende Kompetenzen im Unterricht. Argumentieren in den Fächern Biologie und Religion – ein Beispiel*, in: UTE HARMS/BURKHARD SCHROETER/BARBARA KLÜH (Hrsg.), *Entwicklung kompetenzorientierten Unterrichts in Zusammenarbeit von Forschung*. Münster 2016, 205–223, hier 207.

⁴ Vgl. HANNKEN-ILLJES, *Argumentation*, 21 f.

Topos oder Schlusspräsupposition bezeichnet.⁵ Die drei Bestandteile eines Arguments lassen sich exemplarisch an folgender Grafik veranschaulichen:

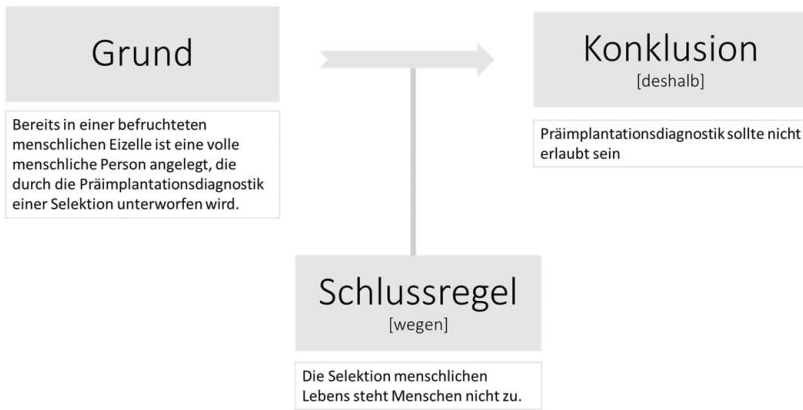


Abb. 7: Das dreiteilige Modell einer Argumentation.

Argumentationen in Alltagssituationen haben selten semantisch vollständige Argumente. Eine oder auch zwei der drei Funktionen bleiben in Alltagsargumentationen oft implizit oder finden keine Berücksichtigung.⁶ Es stellt sich dennoch auch in diesen Kontexten die Frage nach der Plausibilität einer Argumentation, die ein Qualitätsmerkmal von Argumentationen darstellt: Ziel einer Argumentation ist es, die Haltbarkeit und die Relevanz einer Behauptung mittels Bekanntem, Gewusstem oder zwingend Logischem nachzuweisen.⁷ Wenn dies

Eine breit rezipiertes Argumentationsschema stammt von STEPHEN TOULMIN. Neben den drei Grundelementen der Argumentation, welche TOULMIN als »Behauptung« (entspricht der Konklusion), »Daten« (entspricht dem Grund) und der »Schlussregel« (entspricht dem Übergang) bezeichnet, führt er drei weitere Elemente ein, nämlich die »Stützung« der Schlussregeln, womit die Generalisierung und Untermauerung der Gültigkeit der Schlussregel gemeint ist, den »Operator«, der die Relevanz der Daten für die Behauptung konkretisiert, sowie die »Ausnahmebedingungen«, welche die Grenzen der Schlussregel verdeutlichen (vgl. STEPHEN TOULMIN, Der Gebrauch von Argumenten. Weinheim ²1996, 88–98). In Bezug auf diese Argumentationsschemata wendet ELKE VISSER ein, dass sie für den schulischen Kontext zu komplex sind. Sie schlägt stattdessen ein Niveaustufenmodell zur Diagnose von Urteilsbildungsprozessen vor (vgl. 4.1.4).

⁵ HANNKEN-ILLJES, Argumentation, 21 f.

⁶ Vgl. HANNKEN-ILLJES, Argumentation, 22.

⁷ Vgl. MANFRED KIENPOINTNER, Alltagslogik. Struktur und Funktion von Argumentationsmustern. Stuttgart-Bad Cannstatt 1992, 22.

gelingen ist, kann man von Plausibilität einer Argumentation sprechen.⁸ Eine Argumentation ist demnach dann plausibel, wenn »der Weg von der Prämisse zur Konklusion einsichtig und nachvollziehbar« dargestellt ist.⁹ So könnte man beispielsweise im oben dargestellten Argument die Plausibilität der Schussregel anzweifeln, indem man fragt, warum Menschen die Selektion menschlichen Lebens nicht zusteht. Um die Schlussregel plausibler zu gestalten, wären weitere Begründungen zu nennen. So könnte man sich beispielsweise auf die Menschenwürde oder auf eine durch Gott zugesprochene Wertigkeit eines jeden menschlichen Lebens beziehen. Das Ausweisen dieses Bezugs macht es dem Gegenüber möglich, der Aussage zuzustimmen oder sie abzulehnen. Wenn die Gründe nicht ausgewiesen sind, ist eine intersubjektiv nachvollziehbare Zustimmung oder Ablehnung nicht möglich – das Argument bleibt unplausibel.

Neben der Funktion, einen Konsens über eine strittige Frage herzustellen, können Argumentationen auch dazu dienen, Gegensätze zuzuspitzen und gegensätzliche Positionen aufzuzeigen.¹⁰ JOSEF KLEIN betont, dass es auch Gesprächstypen gibt, bei denen der Dissens von vornherein strukturelle Präferenz ist, wie beispielsweise in Fernsehtalkshows oder Streitgesprächen.¹¹ Auch im schulischen Kontext ist das didaktische Ziel des Argumentierens zwischen Schülerinnen und Schülern häufig zunächst, sich die Strittigkeit der Fragestellung zu erarbeiten. KLEIN stellt heraus, dass im aufeinander bezogenen Argumentieren über den Dissens auch ohne einen absehbaren Konsens »die Chance [liegt], dass aus unbegriffenem Dissens begriffener Dissens wird – was bei der Vorbereitung von Entscheidungen häufig notwendige Bedingung für tragfähige Kompromisse ist«.¹² Argumentieren kann folglich auch eine heuristische Funktion als Mittel des Denkens einnehmen – eine Methode, um sich einer Sache klarer zu werden. THERESA SCHWARZKOPF betont darüber hinaus, dass Sprache und die Fähigkeit, Argumente formulieren zu können, sich mit eigenen Konstrukten der Weltdeutung bedingt, da kein Konstrukt ohne passende Sprache gedacht werden kann.¹³ Kommt die Auseinandersetzung mit fremden Deutungen, Konstrukten und Positionen

⁸ Vgl. ebd.

⁹ Vgl. WEIß et al., Fächerübergreifende Kompetenzen, 210.

¹⁰ Vgl. HANNKEN-ILLJES, Argumentation, 25.

¹¹ Vgl. JOSEF KLEIN, Erklären und Argumentieren als interaktive Gesprächskulturen, in: Klaus Brinker, Gerd Antos, Wolfgang Heinemann, Sven F. Sager (Hrsg.), Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung. 2. Halbband. Berlin/New York 2000, 1309–1329, hier 1331.

¹² Vgl. ebd.

¹³ Vgl. SCHWARZKOPF, Vielfältigkeit denken, 26.

In diesem Zusammenhang sei auf LUDWIG WITTGENSTEIN verwiesen, welcher dies pointiert in dem Satz zum Ausdruck brachte: »Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt« (Logisch-Philosophische Abhandlung [*Tractatus logico-philosophicus*] Satz 5.6).

hinzu, so kann sich dadurch das eigene Konstrukt schärfen und die eigene Argumentation (sprachlich) verfeinern.¹⁴ In schriftlichen Abituraufgabenstellungen finden andere Positionen durch Texte und Materialien in die Argumentation Eingang. Wie bereits dargestellt wurde, ist jedoch die Verknüpfung dieser mit den eigenen Thesen Aufgabe der Prüflinge (vgl. Kapitel 3.3.3).

4.1.2 Moralische Urteilsbildung – Lawrence Kohlberg und Georg Lind

Qualitätsmerkmale für Urteilsfähigkeit in Bezug auf kontroverse Fragen wurden in der Psychologie und in der Sozialforschung vor allem unter dem Begriff *moralische Urteilsbildung* erforscht. Der Begriff der *moralischen Urteilsbildung* hat durch die Bezugnahme auf diese Forschungen auch Eintritt in die Fachdidaktiken gefunden.¹⁵

Prägend für die Erforschung moralischer Urteilsbildungsfähigkeit wurde eine Längsschnittstudie von LAWRENCE KOHLBERG, welcher mittels halbstrukturierter Interviews über sogenannte Dilemmageschichten strukturelle Veränderungen von moralischen Urteilen empirisch erhob und in ein Stufenmodell brachte.¹⁶ KOHLBERG zufolge vollzieht sich die moralische Entwicklung eines Menschen in sechs aufeinanderfolgenden Stufen, die er drei Ebenen zuordnet. Die Abfolge erhebt den Anspruch universeller Gültigkeit, wobei auch Unterschiede in Bezug auf die Sozialschicht, das Geschlecht und das Herkunftsland nachweisbar sind.¹⁷ In Hinblick auf Qualitätsmerkmale für moralische Urteilsentscheide ist zentral, dass sich die einzelnen Stufen nicht durch die Entscheidung für oder gegen eine der Handlungsalternativen in moralischen Konfliktsituationen (Dilemmasituation) definieren, sondern durch die Begründungen, mit denen eine Handlungsentscheidung gerechtfertigt wird.¹⁸ Kohlberg beschreibt die Abfolge der Entwicklungsphasen moralischen Urteilens als eine Höherentwicklung, wodurch er eine normative Interpretation seiner Befunde vornimmt.¹⁹ Als Zuordnungs- und durch die wer-

¹⁴ Vgl. SCHWARZKOPF, Vielfältigkeit denken, 26.

¹⁵ Vgl. bspw. CORINNA HÖBLE, Moralische Urteilsfähigkeit. Eine Interventionsstudie zur moralischen Urteilsfähigkeit von Schülern zum Thema Gentechnik. Innsbruck u.a. 2001. MARGRIT MÜLLER, Die Förderung moralischer Urteilskompetenz im Biologieunterricht. Entwicklung eines Analyseinstruments zur Evaluation von Unterrichtskonzepten. Hamburg 2006.

¹⁶ Vgl. LAWRENCE KOHLBERG, Die Psychologie der Moralentwicklung. Hrsg. von WOLFGANG ALTHOF. Frankfurt a. M. 1996.

¹⁷ A. a. O., 58 f.

¹⁸ A. a. O., 26.

¹⁹ Durch die breite Rezeption des Modells und der hohen Zahl an Anschlussforschungen wurde das Modell auch hinsichtlich mehrerer Aspekte kritisiert, wie beispielsweise hin-

tende Struktur der Höherentwicklung auch als Qualitätsmerkmal einer Ebene dienen die Bezugsgrößen, an denen eine Argumentation ausgerichtet wird: Urteilende, die sich auf der untersten Ebene, dem »Prä-moralischen« oder »Präkonventionellen Niveau« befinden, begründen ihre moralischen Urteile strikt aus einer egozentrischen Moralität heraus, die vor allem die unmittelbaren Konsequenzen im Blick hat (Stufe 1 »Orientierung an Strafe und Gehorsam« sowie Stufe 2 »Naiver instrumenteller Hedonismus«).²⁰ Auf der nächst höheren Ebene, dem »Niveau der konventionellen Rollenkonformität«, wird die moralische Entscheidung mittels einer gemeinschaftsbasierten Moralität begründet, die sich am Umfeld des Urteilenden sowie an der gesamten Gesellschaft orientiert (Stufe 3 »Erwartung und Anerkennung« und Stufe 4 »Moral der Aufrechterhaltung von Autorität«).²¹ Auf der höchsten, der häufig als »postkonventionell« bezeichneten Ebene, werden anstelle von gesellschaftlichen Normen oder Gesetzen selbst-akzeptierte moralische Prinzipien zur Begründung eines Urteilsentscheides genannt (Stufe 5 »Moral des Vertrages, der individuellen Rechte und des demokratisch anerkannten Gesetzes/Rechtssystems« und Stufe 6 »Moral der individuellen Gewissensprinzipien«).²² Egozentrischen Argumenten werden hierbei eine geringere Qualität beigemessen, als gemeinschafts- oder gesellschaftsfokussierten Argumentationen. Diese wiederum haben eine geringere Wertigkeit als universalistische Begründungen. Damit wird auch der bei der Argumentation des Urteilsentscheids berücksichtigte Betroffenenkreis mit zunehmender Stufe größer: Während die niedrigen Stufen nur die eigene Person und später die *peer group* im Blick haben, werden später auch gesellschaftliche und auf der höchsten Stufe sogar globale Auswirkungen bedacht.

Die Grundannahme aufgreifend, dass nicht der Urteilsentscheid an sich die Qualität des Urteils ausmacht, sondern die Ausrichtung der angeführten Argumente, entwickelte GEORG LIND einen sogenannten »Moralisches-Urteil-Test«

sichtlich seiner linearen Absolutheit, des Anspruchs auf domänenübergreifende Gültigkeit sowie die Vernachlässigung der emotionalen Dimension des Urteilsbildungsprozesses. Einen Überblick über die Kritik bieten HÜBNER, Philosophische Ethik, 26 f. sowie FRITZ OSER/HORST BIEDERMANN, Art. Moralentwicklung und Moralförderung, in: Handwörterbuch Pädagogische Psychologie 2018, 559–568, hier 560.

²⁰ KOHLBERG, Psychologie der Moralentwicklung, 26.

Vgl. zur Theorie KOHLBERGS auch u. a.: GERHARD BÜTTNER/VEIT-JAKOBUS DIETRICH, Entwicklungspsychologie in der Religionspädagogik. Göttingen 2016, 17–21, sowie HÜBNER, Philosophische Ethik, 26–30.

²¹ Vgl. KOHLBERG, Psychologie der Moralentwicklung, 26.

²² Vgl. ebd.

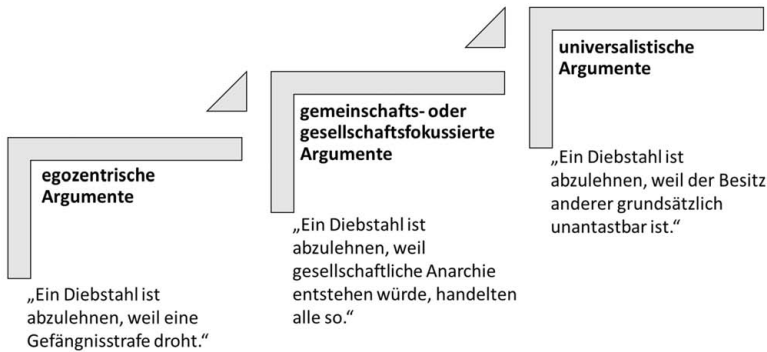


Abb. 8: Die der KOHLBERG'schen Stufentheorie zugrunde liegende Gewichtung von Argumenten am Beispiel der Fragestellung, ob Diebstahl in prekären Situationen legitim sei (angelehnt an das sog. Heinz-Dilemma).

(MUT), welchen er seit 2015 angepasst an den zu untersuchenden Gegenstand »moralische Kompetenz« als »Moralische-Kompetenz-Test« (MKT) bezeichnet.²³

Er unterscheidet zwischen affektiven und kognitiv-strukturellen Aspekten moralischen (Urteils-)Verhaltens. Dabei beschreiben die affektiven Aspekte moralische Leitvorstellungen und Werthaltungen, die einer Person als Orientierung für ihr moralisches Urteil dienen (die Inhalte moralischen Urteilsverhaltens). Die kognitiv-strukturellen Aspekte benennen in Anlehnung an LAWRENCE KOHLBERG die Fähigkeit, »normative Orientierungen integriert und differenziert in konkreten Situationen im Urteilsverhalten zur Geltung zu bringen« (die Struktur moralischen Urteilsverhaltens).²⁴

In LINDS empirischer Forschung, die methodisch mit der Bewertung vorgegebener Argumente zur Begründung von Urteilsentscheiden in Dilemmasituationen arbeitet, wird die Prüfung der moralischen Qualität meinungsunabhängiger Argumente zum Qualitätsmerkmal. LIND schreibt:

»Soweit der Befragte diese [die im MKT vorgegebenen, Anm. KM] Argumente nach moralischen Kategorien differenziert und über die Unterscheidung nach der Meinungskonformität hinweg konsistent beurteilt, gibt er uns zu erkennen, dass er sich in seinem Urteilsverhalten an der moralisch unterschiedlichen Qualität der Argumente orientiert.«²⁵

²³ Vgl. GEORG LIND, Inhalt und Struktur des moralischen Urteilens. Theoretische, methodologische und empirische Untersuchungen zur Moral- und Demokratiekompetenz bei Studierenden. Konstanz ²2000.

²⁴ Vgl. a. a. O., 73 f.

²⁵ A. a. O., 81.

LIND definiert die moralische Kompetenz somit im Wesentlichen an der »moralischen Strukturiertheit«, worunter er die Orientierung an moralischen Kriterien unabhängig der Meinungskonformität versteht. Er bestimmt drei dispositionelle Determinanten moralischen Urteilens: Erstens die »Meinungskonformität«, welche aussagt, ob der Proband ein Argument dahingehend beurteilt, ob es zu seiner persönlichen Meinung passt; zweitens das »Dilemma«, welches beschreibt, dass die Person das Argument der Dilemmasituation angepasst beurteilt sowie drittens »Stufe«, wodurch erfasst wird, inwieweit die Bewertung der Qualität der Argumente den von KOHLBERG erforschten Stufen entsprechen.²⁶ Mit der letzten Determinante ist die Setzung verbunden, dass man KOHLBERG folgend die inhaltliche Qualität von Argumenten objektiv bestimmen kann.

In Weiterentwicklung des MKT entwirft LIND ein *Niveaustufenmodell moralischer Kompetenz*, welches in folgender Tabelle dargestellt wird.²⁷

Niveau I	Vertreten der Meinung ohne Interesse an Argumenten
Niveau II	Hinwendung zu ausschließlich meinungskonformen Argumenten
Niveau III	Auseinandersetzung mit Pro- und Contra-Argumenten, wobei nur den Pro-Argumenten ein Wert beigemessen wird
Niveau IV	Pro- und Contra-Argumente werden nach ihrer moralischen Qualität beurteilt

Tab. 2: Niveaustufenmodell moralischer Kompetenz nach GEORG LIND. Eigene Tabelle nach LIND, *Gewissen lernen?*, 101–111.

Zusammenfassend können aus diesem Kapitel die Merkmale einer Ausrichtung an Argumenten, die einen möglichst großen Personenkreis im Blick haben (KOHLBERG) sowie die Prüfung moralischer Qualität unabhängig der eigenen Position (LIND) als Qualitätsmerkmale für Urteilsfähigkeit festgehalten werden.

4.1.3 Teilkompetenzen ethisch-moralischer Kompetenzniveaus – Das Projekt ETiK

Neben vielen – vor allem auch in Bezug auf das Forschungsdesign – von der Theorie KOHLBERGS geprägten Studien soll mit dem *Berliner Ansatz zur Konstruktion und Erhebung ethisch-moralischer Kompetenzniveaus im öffentlichen Erziehungs- und Bildungssystem* ein Projekt vorgestellt werden, welches nicht

²⁶ A. a. O., 76 und 81.

²⁷ Vgl. GEORG LIND, *Gewissen lernen? Zur Konstanzer Methode der Dilemmadiskussion*, in: ANTON A. BUCHER (Hrsg.), *Moral, Religion, Politik: Psychologisch-pädagogische Zugänge*. Wien 2007, 101–111, hier 106.

unmittelbar in dieser Traditionslinie steht. In dem unter dem Akronym *ETiK*²⁸ stehenden Forschungsprojekt wurde unter Einbezug bildungstheoretischer und fachdidaktischer Konzepte ein Testinstrument entwickelt, welches eine standardisierte und nach Niveaustufen differenzierte Beschreibung ethisch-moralischer Kompetenz ermöglichen soll.²⁹ Das Ziel des Projekts ist es, ein »nützliches Diagnoseinstrument für die fachdidaktische Arbeit und Unterrichtsentwicklung« für das Fach Ethik sowie für die vergleichbaren Fächer zu entwickeln.³⁰

DIETRICH BENNER und ROUMIANA NIKOLOVA definieren drei domänenspezifische Teilkompetenzen einer durch Ethik-Unterricht und in vergleichbaren Fächern »zu entwickelnden und zu stärkenden moralischen Kompetenz«³¹: 1. »Ethisch-moralische Grundkenntnisse«, 2. »Ethisch-moralische Urteilskompetenz« sowie 3. »Ethisch-moralische Handlungsentwurfskompetenz«, wobei unter letzterer nicht die Herausbildung eines tugendhaften Charakters intendiert wird, sondern eine Kompetenz beschrieben werden soll, die »moralische Entscheidungen begründen und moralische Handlungen mit anderen abstimmen zu können erlaubt.«³² Es handelt sich bei der Unterscheidung um idealtypische Dimensionen, welche in einem wechselseitigen Verweiszusammenhang stehen und daher nicht hierarchisch gedacht werden.³³ Das Modell wurde anhand einer Validierungsstudie bestätigt und als empirisch gehaltvoll erachtet.³⁴

²⁸ ETiK steht für »Entwicklung eines Testinstruments zu einer didaktisch und bildungstheoretisch ausgewiesenen Erfassung moralischer Kompetenzen, bezogen auf den Ethik-Unterricht an öffentlichen Schulen«.

²⁹ Vgl. STANISLAV IVANOV, Die Validierung des ETiK-Instruments, in: DIETRICH BENNER/ROUMIANA NIKOLOVA (Hrsg.), *Ethisch-moralische Kompetenz als Teil öffentlicher Bildung*. Paderborn 2016, 101–129, hier 129.

³⁰ Vgl. ebd.

³¹ Vgl. DIETRICH BENNER/ROUMIANA NIKOLOVA, Der Berliner Ansatz zur Konstruktion und Erhebung ethisch-moralischer Kompetenzniveaus im öffentlichen Erziehungs- und Bildungssystem, in: DIES. (Hrsg.), *Ethisch-moralische Kompetenz als Teil öffentlicher Bildung*. Paderborn 2016, 13–68, hier 32.

³² Vgl. a. a. O., 32 f.

Moralische Kompetenz wird im Projekt ETiK als eine kognitive Kompetenz definiert. BENNER et al. schreiben an anderer Stelle: »Die Testaufgaben werden ferner so formuliert, dass sie nicht Einstellungen oder Meinungen erheben, sondern domänenspezifisch identifizierbare Kompetenzen erfassen – und anders als die Kompetenzdefinition von Weinert, welche motivationale und volitionale Momente von Kompetenzen benennt (vgl. BMBF [Hrsg.], *Bildungsstandards*, 24) – eine kognitive Fokussierung auf den kognitiven Kern eines jeweiligen Faches aufweisen.« DIETRICH BENNER/SHAMSI DEGHANI/ROUMIANA NIKOLOVA/JOANNA SCHARREL/ROLF SCHIEDER/HENNING SCHLUB/THOMAS WEIB/JOACHIM WILLEMS, Modellierung und Testung religiöser und ethischer Kompetenzen im Interesse ihrer Vergleichbarkeit, in: *Zeitschrift für Pädagogik und Theologie* 2 (2010). 165–174, hier 167.

³³ Vgl. BENNER/NIKOLOVA, *Ethisch-moralische Kompetenzniveaus*, 68.

Für alle drei Teilkompetenzen moralischer Kompetenz wurden mittels auf einer Rasch-Skala erkennbarer Schwellen, »die qualitative Sprünge in den schwierigkeitsgenerierenden Anforderungsmerkmalen der Aufgaben markieren«³⁵, Anforderungsniveaus differenziert. Im Bereich ethisch-moralische Urteilskompetenz wurden folgende Anforderungsniveaus erhoben³⁶:

Niveau I	Die Schülerinnen und Schüler können »[...] moralische Fragen und Problemstellungen in erfahrungsnahen Kontexten angemessen beurteilen und nach schon erworbenen konventionellen moralischen Argumentationsmustern und Denkweisen beurteilen.«
Niveau II	Die Schülerinnen und Schüler können »[...] moralische Konfliktfälle auch jenseits der ihnen vertrauten Erfahrungswelt reflektieren und bearbeiten.«
Niveau III	Die Schülerinnen und Schüler können »[...] moralische Problemstellungen aus partiell divergierenden Moralen reflektieren und beurteilen.«
Niveau IV	Die Schülerinnen und Schüler können »[...] gesellschaftliche Werte und Normen unter verschiedenen ethischen Fragestellungen und Betrachtungsweisen reflektieren sowie auf ihre Gültigkeit hin befragen.«

Tab. 3: Die im Projekt ETiK identifizierten Niveaustufen ethisch-moralischer Urteilskompetenz. Eigene Tabelle nach BENNER/NIKOLOVA, Anforderungsniveaus der ethisch-moralischen Teilkompetenzen, 141–144.³⁷

Durch die Niveaustufen wird als schwierigkeitskonstituierendes Merkmal für ethisch-moralische Urteilsbildung die Erfahrungsferne der Fragestellung definiert. Ähnlich wie im Modell LINDS taucht auch in diesem Modell unter Niveaustufe drei die Multiperspektivität als Qualitätsmerkmal auf. Als höchste Qualitätsmerkmale werden das Urteilen auch in lebensweltlich nicht vertrauten Problemlagen sowie das Prüfen gesellschaftlicher Werte und Normen benannt.

³⁴ Vgl. ebd.

³⁵ DIETRICH BENNER/STANISLAV IVANOV/ROUMIANA NIKOLOVA, Die Anforderungsniveaus der ethisch-moralischen Teilkompetenzen, in: DIETRICH BENNER/ROUMIANA NIKOLOVA (Hrsg.), *Ethisch-moralische Kompetenz als Teil öffentlicher Bildung*. Paderborn 2016, 131–146, hier 135.

»Solche Merkmalsunterschiede können z.B. in den zum Lösen der Aufgaben notwendigen fachspezifischen Kenntnissen oder in der Komplexität der zur Lösung führenden kognitiven Prozesse liegen. Die Schwellen können mit Hilfe sogenannter »Schwellenitems« identifiziert werden, die einen qualitativen Anforderungssprung zum nächst höher liegenden Niveau markieren« (a. a. O.).

³⁶ Vgl. a. a. O., 141–144. Der Bestimmung der Anforderungsniveaus lagen 34 Items zugrunde.

³⁷ Vgl. ebd.

4.1.4 Forschungen zu Bewertungskompetenz aus der Biologiedidaktik

Weitere Forschungen zu Struktur- und Niveaustufenmodellen von Urteilskompetenz sind in den vergangenen Jahren vor allem in der Biologiedidaktik vorgenommen worden. Anlass dazu war die Verankerung von Urteilsbildungsprozessen durch den eigenständigen Kompetenzbereich »Bewertung« in den Bildungsstandards für das Fach Biologie im Jahr 2004.³⁸ Vor dem Hintergrund fachdidaktischer Forschungen aus anderen Fächern wurde der Frage nachgegangen, wie sich der neue Kompetenzbereich für das Fach Biologie beschreiben, fördern und diagnostizieren lässt.³⁹ NEELE ALFS, KERSTIN HEUSINGER VON WALDEGGE und CORINNA HÖBLE stellen ein *Strukturmodell ethischer Bewertungskompetenz* vor, welches Strukturmodelle zur Urteilskompetenz aus anderen Fachdidaktiken aufgreift, sie für die Biologiedidaktik ausdifferenziert und den Anspruch erhebt, sieben Teilkompetenzen zu benennen, welche Bewertungskompetenz vollständig abbilden:⁴⁰

1. Wahrnehmung und Bewusstmachen der eigenen Einstellungen
mit dem Ziel der Distanzierung zu und dem Bewusstmachen von eigenen Einstellungen.
2. Wahrnehmen und Bewusstmachen moralischer Relevanz
mit dem Ziel, den moralischen Gehalt der Fragestellung benennen und begründen zu können. Dabei sollen die zur Disposition stehenden Werte definiert werden.
3. Beurteilen
mit dem Ziel, Anti- und Sympathien aufgrund von Werten reflektieren zu können.

³⁸ Vgl. STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER (Hrsg.), Beschlüsse der Kultusministerkonferenz. Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004. München/Neuwied 2005, 12 und 15.

Durch die Benennung der neu eingeführten Kompetenz als »Bewertungskompetenz« in den Bildungsstandards der KMK sind auch die fachdidaktischen Forschungen im Fach Biologie an diesen Begriff angelehnt. Inhaltlich geht es um die Bewertung kontroverser, bioethischer Fragestellungen, so dass »Bewertungskompetenz« mit dem in vielen anderen Fachdidaktiken geläufigeren Begriff der »Urteilskompetenz« bzw. »-fähigkeit« gleichzusetzen ist.

³⁹ Vgl. zusammenfassend NEELE ALFS/KERSTIN HEUSINGER VON WALDEGGE/CORINNA HÖBLE, Bewertungsprozesse verstehen und diagnostizieren, in: Zeitschrift für interpretative Schul- und Unterrichtsforschung 1 (2012) 1, 83–112, hier 83 f. Zur Biodidaktischen Forschung über Niveaunkonkretisierungen der Urteilskompetenz siehe auch KATJA REITSCHERT/JÜRGEN LANGLET/CORINNA HÖBLE/NICOLA MITTELSTEN SCHEID/KIRSTEN SCHLÜTER, Dimensionen Ethischer Urteilskompetenz. Dimensionierung und Niveaunkonkretisierung, in: Der mathematische und naturwissenschaftliche Unterricht 60 (2007) 1, 43–51.

⁴⁰ Vgl. ALFS et al., Bewertungsprozesse, 87–91.

4. Folgenreflexion
mit dem Ziel, die Folgedimensionen eines Urteils antizipieren zu können (kurzfristig/langfristig, real/weniger real, mittelbar/unmittelbar betroffene Personen/gesellschaftliche Veränderungen).
5. Perspektivwechsel
mit dem Ziel, von der Gebundenheit einer Perspektive zu abstrahieren und eine allgemeine gesellschaftliche Perspektive einzunehmen.
6. Argumentieren
mit dem Ziel, schlüssige Argumente zu formulieren, die der formalen Logik folgen.
7. Ethisches Basiswissen
als eine Kategorie, die alle anderen Teilkompetenzen stützt. Sie umfasst das Nutzen und Reflektieren ethischer Begriffe, Methoden und Konzepte.

Dieses Strukturmodell nennt wesentliche Teilaspekte einer Urteilsbildung, aus deren Vorhandensein sich die Qualität eines Urteilsentscheides ergibt. Mehrere Aspekte der in den vorangegangenen Abschnitten benannten Qualitätsmerkmale lassen sich in dem Modell wiederfinden, wie beispielsweise die Multiperspektivität, die in der Teilkompetenz »Perspektivwechsel« enthalten ist oder die Argumentationsfähigkeit. Es formuliert jedoch auch noch nicht genannte Aspekte, beispielsweise, dass das Wahrnehmen der moralischen Relevanz durch das Benennen der zur Disposition stehenden Werte gefasst werden kann, oder die Folgenreflexion eines Urteilsentscheides als Qualitätsmerkmal. Das Modell trifft Aussagen über die Struktur der Kompetenz, jedoch nicht über Niveaustufungen (vgl. Kapitel 3.3.1). Daher soll ein ebenfalls aus der Biologiedidaktik stammendes Modell den ersten Teil des Kapitels abschließen, welches Bewertungskriterien für Urteilskompetenz in drei Niveaustufenmodellen fasst.

Zur Bestimmung dieser untersuchte ELKE VISSER die ethische Bewertungskompetenz von Schülerinnen und Schülern des elften Schuljahrgangs im Biologieunterricht anhand zweier Fragestellungen zur Legalisierung der Eizellspende und zur Klonierung.⁴¹ Im Gegensatz zum vorangegangenen Modell unterscheidet sie lediglich drei Diagnosebereiche in Bezug auf die Urteilsbegründung, für welche sie jeweils drei bis vier Niveaustufen auf Grundlage der fachdidaktischen Forschungsdiskurse über Urteilsfähigkeit benennt. Die von ihr benannten Diagnosebereiche umfassen deskriptive Aspekte, normative Aspekte und argumentative Kompetenz. Die Niveaustufen sind nicht wie im oben dargestellten ETiK-Projekt induktiv erhoben, sondern deduktiv auf Grundlage der Forschungsliteratur von VISSER gesetzt und im Anschluss hinsichtlich ihres diagnostischen Nutzens empirisch geprüft.

Im Hinblick auf die Struktur von Argumentationsmustern kommt sie zu dem Ergebnis, dass im Rahmen eines Diskurses sehr selten vollständige Argumen-

⁴¹ Vgl. VISSER, Diagnose der Bewertungskompetenz.

tationen im Sinne der Argumentationstheorien Verwendung finden und schließt daraus, dass starre Argumentationsmuster, wie beispielsweise der praktische Syllogismus oder das Schema nach TOULMIN für eine Diagnose der Argumentationsstruktur zu anspruchsvoll bzw. »überdimensionalisiert« sind (zur Argumentationstheorie vgl. Kapitel 4.1.1).⁴² Sie schlägt daher ein Modell mit fünf grundlegenden Argumentationselementen für die Diagnose von Begründungen vor:

»Zur Formulierung einer Stellungnahme (1) treten rechtfertigende Elemente. Diese lassen sich in das direkt stützende funktionale Element (2) (X sollte getan/unterlassen werden, *weil...*) und das konditionale Argument (3) einteilen, das einschränkende Bedingungen anführt (X sollte getan/unterlassen werden, *wenn...*). Hinzu kommen die Nennung von Gegenargumenten (4) sowie die Abwägung der Argumente.«⁴³

Somit sind folgende Merkmale für die Plausibilität einer Argumentation von Bedeutung:

- »Formulierung einer Stellungnahme
- Nennung funktionaler Argumente
- Nennung von Gegenargumenten
- Abwägung der Argumente unterschiedlicher Positionen
- Nennung eingeschränkter Bedingungen: Nennung konditionaler Argumente (Auseinandersetzung mit potenziellen Gegenargumenten).«⁴⁴

VISSER unterscheidet in Bezug auf die Argumentation die im Folgenden tabellarisch dargestellten Niveaustufen:

Niveau I	Es werden keine rechtfertigenden Argumente genannt (weder funktional noch konditional).
Niveau II	Es wird mindestens ein rechtfertigendes Argument genannt.
Niveau III	Es werden rechtfertigende Argumente für mindestens zwei unterscheidbare Positionen benannt.
Niveau IV	Die Argumente für unterscheidbare Positionen werden abgewogen und gewichtet.

Tab. 4: Niveaustufen für die Teilkompetenz Argumentieren. Eigene Tabelle nach VISSER, Diagnose der Bewertungskompetenz, 280.

⁴² Vgl. VISSER/HÖBLE, Bioethisch argumentieren, 189.

⁴³ A. a. O., 190. Hervorhebungen im Original.

⁴⁴ VISSER, Diagnose der Bewertungskompetenz, 292 f.

Als Qualitätsmerkmal fungiert also zunächst die Begründung eines Urteils an sich. Hinzu tritt die Multiperspektivität und ähnlich wie bei LIND die Gewichtung der Argumente unterscheidbarer Positionen. Um etwas über die Qualität der jeweiligen Argumente auszusagen, bedarf es jedoch einer weiteren Analyse. VISSER schlägt vor, auf Grundlage des praktischen Syllogismus, die Bezugnahme von deskriptiven und normativen Prämissen zu diagnostizieren. Im Bereich der deskriptiven Prämissen stellt VISSER vor allem die Beschreibung der Folgen einer Handlung sowie die Berücksichtigung der Perspektiven der Beteiligten und Betroffenen als leitende Qualitätskriterien heraus. Sie benennt folgende Niveaustufen:

Niveau I	Es werden keine Folgen oder nur Folgen für unmittelbar Beteiligte genannt.
Niveau II	Es werden Folgen anderer als die der unmittelbar Beteiligten benannt. Die Folgen gehen über die unmittelbar angestrebten Handlungsergebnisse bzw. über die ausdrücklich im Material genannten Folgen hinaus.
Niveau III	Die Folgenreflexion berücksichtigt Wirkungen auf gesellschaftlicher Ebene.

Tab. 5: Niveaustufen der deskriptiven Aspekte. Eigene Tabelle nach VISSER: Diagnose der Bewertungskompetenz, 289.

Im Bereich der normativen Prämissen stellt sich VISSER zufolge die Frage nach den der Argumentation zugrunde liegenden Werten und Normen. Als Qualitätsmerkmale werden die Verwendung von Werten und Normen sowie deren Offenlegung auf einer Metaebene benannt. In diesem Zusammenhang definiert VISSER folgende vier Niveaustufen:

Niveau I	Es werden keine normativen Elemente wie Wertungen, Werte und Normen verwendet.
Niveau II	Es werden Normen und Werte verwendet.
Niveau III	Normen und Werte werden explizit genannt.
Niveau IV	Wertungen, Normen und Werte werden ausdrücklich als solche bezeichnet.

Tab. 6: Niveaustufen normative Aspekte. Eigene Tabelle nach VISSER, Diagnose der Bewertungskompetenz, 278.

Das Modell ist zugunsten einer Nutzung im Unterrichtsalltag bewusst einfach gehalten – die daraus folgende geringere Präzision wird von VISSER diskutiert.⁴⁵ Sie hält fest, dass die einzelnen Niveaustufen im Hinblick auf die konkrete Fragestellung ergänzungsbedürftig und -fähig sind. Dennoch geben sie einen Orientierungsrahmen für eine analytische Bewertung.

⁴⁵ Vgl. a. a. O., 287 sowie 281–283.

4.2 Religionspädagogischer Forschungsstand

Die religionspädagogische Forschung stellt Urteilsfähigkeit in den größeren Kontext des ethischen Lernens. Das Nachdenken über das Konstrukt Urteilsfähigkeit nimmt daher in rechtlichen Überlegungen über die Verortung des Staates zu Bekenntnisfragen und in einer Auseinandersetzung mit der Rolle konfessioneller Bestimmtheit in ethischen und urteilsbildenden Lernprozessen seinen Ausgangspunkt. Deshalb werden diese – bereits in Kapitel 2.3 skizzierte – Aspekte zunächst noch einmal aufgegriffen, um von ihnen ausgehend religionspädagogische Konzeptionen und Modelle zur Urteilsfähigkeit vorzustellen. Ziel der Darstellung ist, die in den Konzepten explizit und implizit enthaltenen, gegebenenfalls spezifisch religionspädagogischen Qualitätsmerkmale herauszuarbeiten.

4.2.1 Religiös-weltanschauliche Neutralität des Staates in Urteilsfragen? Eine rechtliche Perspektive

Der deutsche Staat hat sich zur religiös-weltanschaulichen Neutralität, dem sogenannten »Prinzip der Nicht-Identifikation« verpflichtet, welches durch mehrere Artikel im Grundgesetz verankert ist (Art. 4 Abs. 1, Art. 3 Abs. 3 sowie Art. 33 Abs. 3). HEINER BIELEFELDT weist darauf hin, dass diese religiös-weltanschauliche Neutralität nicht mit einer generellen Wertneutralität gleichzusetzen ist, sondern aus der Achtung vor der Religionsfreiheit der Menschen resultiert, die durch das Grundgesetz den Status eines hohen Verfassungswertes, nämlich eines allgemeinen Menschenrechts zugesprochen bekommt.⁴⁶ Der Staat unterliegt folglich keiner »normativen Bindungslosigkeit« (BIELEFELDT), sondern bindet sich ethisch und rechtlich an Werte, wie etwa den der Religionsfreiheit oder der Menschenwürde.⁴⁷ Er legt sich dabei selbst das Verbot auf, sich mit einer bestimmten Religion zu identifizieren, um andere Religionen und auch konfessionslose Menschen nicht zu benachteiligen. Wenn der Staat sich verfassungsrechtlich zu Werten bekennt, dann also nur so weit, als dass er die Begründung für diese Werte offenlässt und sich somit keiner den Wert begründenden Letztdeutung anschließt, wie sie beispielsweise die Religionen oder der Humanismus bieten.⁴⁸ Dadurch wird sichergestellt, dass »die in Artikel 1 Grundgesetz formulierte rechtsethische Bindung an die obersten Verfassungswerte nicht zu

⁴⁶ Vgl. HEINER BIELEFELDT, Muslimische Minderheiten im säkularen Rechtsstaat, in: GERHARD KRIUP/WOLFGANG VÖGELE (Hrsg.), Schatten der Differenz. Das Paradigma der Anerkennung und die Realität gesellschaftlicher Konflikte. Hamburg 2006, 69–90, hier 71.

⁴⁷ Vgl. BIELEFELDT, Muslimische Minderheiten im säkularen Rechtsstaat, 71.

⁴⁸ Vgl. a. a. O., 73.

einer religiösen oder weltanschaulichen Fundierung des Staates gerät«⁴⁹. Analog gilt auch, dass in der unterrichtlichen Wertebildung die Übernahme von wertebegründenden Letztdeutungen nicht von den Schülerinnen und Schülern gefordert werden darf. »Unterricht, der Informationen über ethische Konzepte vermittelt und dabei wertend Position bezieht, ist so lange verfassungsrechtlich unbedenklich, als nicht indoktrinierend eine einzelne Sinndeutung des menschlichen Lebens den Schülern auferlegt wird (Prinzip der Nichtidentifikation, Verbot des ›Oktroi‹).«⁵⁰

Der Religionsunterricht – auch wenn er sich entschieden von einer Indoktrination distanziert – nimmt durch seine verfassungsrechtlich verankerte konfessionelle Bindung eine Sonderstellung ein. Nach Art. 7.3 wird er »in Übereinstimmung mit den Grundsätzen der Religionsgemeinschaften« erteilt und ist somit der einzige Unterricht, in welchem »dem Beweise unzugängliche Glaubenssätze akzeptanzheischend vermittelt werden«⁵¹ dürfen. Dementsprechend gilt für die Wertebildung im Religionsunterricht, dass sie sich an den Werten der jeweiligen Religionsgemeinschaft ausrichtet. Aus rechtlicher Sicht steht dem Religionsunterricht also zu, nicht beweisbare Glaubenssätze zu vermitteln. Das Bundesverfassungsgericht hat dies im Februar 1987 nochmals in aller Deutlichkeit herausgestellt:

»Er ist keine überkonfessionelle vergleichende Betrachtung religiöser Lehren, nicht bloße Morallehre, Sittenunterricht, historisierende und relativierende Religionskunde, Religions- oder Bibelgeschichte. Sein Gegenstand ist vielmehr der Bekenntnisinhalt, nämlich die Glaubenssätze der jeweiligen Religionsgemeinschaft. Diese als bestehende Wahrheiten zu vermitteln ist seine Aufgabe. (BVerfGE 74,244,252).«⁵²

In der religionsdidaktischen Debatte darüber, was im evangelischen Religionsunterricht vermittelt werden soll, ist man diesbezüglich jedoch zurückhaltender. Als Ziel wird häufig eine kritische, distanzierte Reflexion des eigenen Glaubens gefordert, anstatt eine Beheimatung in der eigenen Religion und Tradition anzustreben, wie es im katholischen Religionsunterricht eher der Fall ist. Dass das spezifisch evangelische Bekenntnis sowie die damit verbundenen Letztbegründungen und Wahrheitsansprüche eine zentrale Rolle im Religionsunterricht und damit auch in den dort stattfindenden Urteilsprozessen haben dürfen, können

⁴⁹ Vgl. ebd.

⁵⁰ PHILIP KUNIG, Rechtsfragen ethischer und religiöser Erziehung in der Schule, in: GOTTFRIED ADAM/FRIEDRICH SCHWEITZER (Hrsg.), *Ethisch erziehen in der Schule*. Göttingen 1996, 301–312, hier 306.

⁵¹ A. a. O., 307.

⁵² Zitiert nach UTA HILDEBRANDT, *Das Grundrecht auf Religionsunterricht*, Tübingen 2000, 61, Anm. 95.

und sollten, erweist sich angesichts der Rechtstexte und der Profilierung des Faches als unbedingt notwendig.

Ein gängiger Ansatz, die konfessionelle Bestimmtheit und die freien Urteilsbildung in ein Verhältnis zu setzen, besteht in der Auffassung, dass die Auseinandersetzung mit christlichen Weltansichten die Ausdifferenzierung und die kritische Reflexion der eigenen Weltansicht fördert, ohne sie zu determinieren. In dieser Weise argumentiert auch INGRID SCHOBERTH, die die christliche Bestimmtheit als Orientierung benennt, während die Urteilsbildung der Schülerinnen und Schüler in Bezug auf die Positionierung offen ist und bleiben muss. In folgendem Absatz wird SCHOBERTHS Begründung deutlich:

»Die Lebensform christlicher Religion steckt also den Rahmen ab, auf den religiöse Bildungsprozesse bezogen sind und sie bleibt der bestimmende Bezugspunkt der Auseinandersetzungen, die letztlich immer eine offene Gestalt haben müssen: Religiöse Bildungsprozesse vollziehen sich in Offenheit und Bestimmtheit bezogen auf christliche Religion und schaffen so einen *Diskurskontext*, in dem sich Kinder und Jugendliche bewegen lernen und zur Auseinandersetzung mit vielfältigen Perspektiven angeleitet werden. [...] Die Orientierungen und Perspektiven der christlichen Lebensform werden im Lernen ebenso zur Sprache gebracht, wie eine mögliche Abgrenzung und Verweigerung dieser Lebensform gegenüber, weil die Kinder und Jugendlichen sie nicht (oder noch nicht/nicht mehr) teilen wollen. [...] In dieser Hinsicht bleibt auch das Urteilen lernen *missionarisch* in dem Sinne, dass den Kindern und Jugendlichen christliche Religion mit ihren Orientierungen und starken Bewertungen gezeigt wird (Bestimmtheit), ohne damit freilich einzufordern, dass Kinder und Jugendliche diese Perspektiven und Orientierungen teilen müssen (Offenheit).«⁵³

Inwieweit Urteilsbildung in diesem Sinne tatsächlich als missionarisch verstanden werden kann und darf, ist fraglich und wird in der Religionspädagogik kontrovers diskutiert, wie bereits in Kapitel 2.3 deutlich wurde.

RUDOLF ENGLERT hält die Ausbildung moralischer Grundhaltungen für eine Etappe im ethischen Lernen, die der Urteilsbildung vorausgehen sollte.⁵⁴ Er plädiert dafür, Urteilsbildung als die komplexeste Stufe ethischen Lernens zu

⁵³ INGRID SCHOBERTH, Urteilen lernen. Einleitende Reflexionen, Perspektiven und Orientierungen in religionspädagogischer Perspektive, in: DIES. (Hrsg.), Urteilen lernen – Grundlegung und Kontexte ethischer Urteilsbildung. Göttingen/Oakville 2012, 25–40, hier 32 f. Hervorhebungen wie im Original.

⁵⁴ Vgl. RUDOLF ENGLERT, Die verschiedenen Komponenten ethischen Lernens und ihr Zusammenspiel. Überlegungen zu einem Gesamtprogramm ethischer Bildung, in: RUDOLF ENGLERT/HELGA KOHLER-SPIEGEL/ELISABETH NAURATH/BERND SCHRÖDER/FRIEDRICH SCHWEITZER (Hrsg.), Ethisches Lernen. Neukirchen-Vluyn 2015 [= Jahrbuch der Religionspädagogik 31], 108–118.

verstehen, in der ethisches Lernen jedoch nicht aufgeht, sondern welcher die Entwicklung von Tugenden und die Reflexion von Werten vorangehen sollten, um ein – verschiedene Ansätze einschließendes – *Gesamtkonzept ethischen Lernens* zu verfolgen.⁵⁵

Einen eher rationalistischen Ansatz der Urteilsbildung verfolgt BERND SCHRÖDER, wenn er sich für eine höhere Relevanz der Konfessionalität in Urteilsbildungsprozessen ausspricht. Er schreibt:

»[I]m Lichte der ethischen Überlieferungen von Religionen und Denominationen [kann es] weder inhaltlich noch methodisch dasselbe sein, ob ein ethisches Thema in dieser oder in jener Spielart von Religionsunterricht und/oder Ethikunterricht behandelt wird – und es sollte auch nicht dasselbe sein, wenn ein differenziertes Unterrichtsangebot im Sinne von Art. 7.3 GG genutzt werden und geboten sein soll.«⁵⁶

Die bildungsadministrativen Vorgaben zeichnen jedoch ein anderes Bild: Nach einer exemplarischen Untersuchung von niedersächsischen Kerncurricula des evangelischen, katholischen, jüdischen und islamischen Religionsunterrichts sowie des Faches Werte und Normen (welches in Niedersachsen das Äquivalent zum Ethikunterricht anderer Bundesländern ist) kommt SCHRÖDER zu dem Befund, dass die Kerncurricula »die Unterschiede zwischen den ethischen Traditionen und Gegenwartsdiskursen der unterrichteten Religionen und Denominationen nur unzureichend wider[spiegeln]«. ⁵⁷ Weiter kritisiert er, dass die untersuchten Kerncurricula sowohl Lehrenden als auch den Lernenden wenig Handhabe böten, »sich in der Pluralität ethischer Perspektiven zu orientieren, die Auseinandersetzung zu suchen und das Potential des gemeinsamen Ethos zu würdigen. Insbesondere der je und je konfessionell bestimmte Religionsunterricht der christlichen Konfessionen«, so SCHRÖDER weiter, »versäumt es im Zeichen der Kompetenzorientierung, die je spezifische ethische Perspektive der eigenen Religionsgemeinschaft, aber auch anderer christlicher Denominationen zu profilieren.« ⁵⁸ SCHRÖDER fordert daher, dass die denominationsspezifischen Perspektiven im jeweiligen Religionsunterricht nicht bloß über die Auswahl verschiedener Themen, sondern auch a) durch die Auswahl der Texte über traditionelle oder gegenwärtige Fragestellungen aus der jeweiligen Religionsge-

⁵⁵ Vgl. ebd.

⁵⁶ BERND SCHRÖDER, Was macht es für einen Unterschied, ob ethische Fragen im Ethik- oder evangelischen / katholischen / jüdischen / islamischen Religionsunterricht behandelt werden?, in: RUDOLF ENGLERT/HELGA KOHLER-SPIEGEL/ELISABETH NAURATH/BERND SCHRÖDER/FRIEDRICH SCHWEITZER (Hrsg.), Ethisches Lernen. Neukirchen-Vluyn 2015 [= Jahrbuch der Religionspädagogik 31], 41–63, hier 51.

⁵⁷ Ebd.

⁵⁸ Ebd.

meinschaft respektive Theologie; b) durch ethische Kriterien der Urteilsfindung und -begründung (seien diese explizit oder implizit); und c) durch die Methoden der Urteilsfindung erkennbar werden sollten.⁵⁹ Das von ihm vertretene Konzept ethischer Urteilsbildung ist zudem darauf angelegt, in der Auseinandersetzung mit verschiedenen religions- und denominationsspezifischen Perspektiven die eigene Sicht zu hinterfragen.

Die von SCHRÖDER geforderte Stärkung der konfessionellen Profile kann jedoch auch kritisch hinsichtlich ihrer Konsequenzen betrachtet werden. So ist zu fragen, welchen Ertrag eine Stärkung der unterschiedlichen ethischen Traditionen für das gesamtgesellschaftliche Zusammenleben hat, welches einen Aushandlungsprozess gemeinsamer Werte bedarf.

4.2.2 Modellbildungen ethischer Urteilsbildung

Neben der Diskussion über die Rolle der konfessionellen Bindung des Religionsunterrichts in Urteilsbildungsprozessen gibt es religionspädagogische Forschungen, die den Urteilsbildungsprozess durch Modellbildungen genauer zu beschreiben versuchen und dadurch auch wesentliche Qualitätsmerkmale einer Urteilsbildung herausstellen. Anders als bei den im Vorangegangenen beschriebenen Bestimmungen von Urteilsfähigkeit versuchen die im Folgenden betrachteten Urteilsbildungsmodelle den Urteilsbildungsprozess in einer zirkulär verstandenen Handlungsabfolge zu beschreiben. Wesentliche Schritte sind dabei – nach der inhaltlichen und ethischen Erfassung der Fragestellung – die bereits als Qualitätsmerkmale benannte Betrachtung verschiedener Handlungsfolgen und Perspektiven sowie die darauf basierende Formulierung von Argumenten.⁶⁰ Nach der Prüfung der den Argumenten zugrunde liegenden Nor-

⁵⁹ Vgl. a. a. O., 51 f.

⁶⁰ Stufenmodelle dieser Art lassen sich sowohl in der Religionspädagogik als auch in anderen Fachdidaktiken finden. Vgl. bspw. CHRISTOPH BAUMGARTNER, Umweltethik – Umwelthandeln. Ein Beitrag zur Lösung des Motivationsproblems. Paderborn 2005, 47; ZWERGEL, Kompetenzen und Grenzen ihrer Standardisierung, 99–104; CHRISTIANE MEYER/DIRK FELZMANN, Was zeichnet ein gelungenes ethisches Urteil aus? Ethische Urteilskompetenz im Geographieunterricht unter der Lupe, in: CHRISTIANE MEYER/RODERICH HENRY/GEORG STÖBER (Hrsg.), Geographische Bildung. Kompetenzen in didaktischer Forschung und Schulpraxis. Braunschweig 2011, 130–146; KURT LACH, Planung von zwei Unterrichtsstunden: Der Castor-Transport – »Bürgerkrieg« im Wendland. War das Handeln der Beteiligten gerechtfertigt?, in: GOTTHARD BREIT (Hrsg.), Politische Urteilsbildung. Aufgabe und Wege für den Politikunterricht [= Schriftenreihe Bundeszentrale für politische Bildung, Bd. 344]. Bonn 1997, 159–178; EVA-MARIE ULRICH-RIEDHAMMER/STEFAN APPLIS, Ethisches Argumentieren als Herausforderung, in: Praxis Geographie 3 (2013) 24–29.

men und Werte werden diese abgewogen, was zu einem reflektierten Urteilsentscheid führen soll. Vor allem in der Religionspädagogik viel zitiert ist eine *Ethische Theorie sittlicher Urteilsbildung* von HEINZ EDUARD TÖDT, in der er 1988 ein solches Modell ethischer Urteilsbildung vorlegte.⁶¹ Dieses Modell wurde verschiedentlich weiterentwickelt. Ein neuerer, die vorangegangenen Theorien aufnehmender Entwurf stammt von HANS-RICHARD REUTER, welcher vier Arbeitsschritte benennt, die im Folgenden skizziert werden:⁶²

1. Beschreibung des Kontextes

Die Beschreibung des Kontextes besteht aus a) der »Analyse und Kenntnisnahme der relevanten empirischen Fakten und Rahmenbedingungen«, b) der Kenntnis der rechtlichen Rahmenbedingungen⁶³, c) der Berücksichtigung derer, die von einem ethischen Problem betroffen sind (Interessen, Ansprüche, aus ersteren resultierende Lösungsmöglichkeiten, Machtverhältnisse und Einflussmöglichkeiten) und d) aus der Analyse des gelebten Ethos, problemspezifischer Intuitionen und etablierter Überzeugungen.

2. Identifizierung der ethischen Perspektive(n), Kriterien und Ordnungen

In einem zweiten Arbeitsschritt geht es um die Wahrnehmung des spezifisch ethischen Aspekts des Problems und die Analyse moralischer Pflichten und Normen (höchste Leitkriterien, Güter und Ziele, »gelingendes Leben«). Er beinhaltet a) die Präzisierung von Perspektiven und den damit verbundenen Leitkriterien: normativ (Was sollen wir tun?), evaluativ (Wie wollen wir leben?), konsultativ (Was befähigt dazu, gut handeln zu können?), b) das Ausweisen umfassender, religiöser oder philosophischer Wirklichkeitsannahmen sowie c) das Identifizieren der konkreten handlungsleitenden Orientierungen in Bezug auf die Fragestellung (bereichsspezifische Pflichten und Normen, sozial geteilte Güter und Werte).

⁶¹ Vgl. HEINZ EDUARD TÖDT, *Perspektiven theologischer Ethik*. München 1988.

⁶² Vgl. REUTER, *Grundlagen und Methodik der Ethik*, 112–116.

⁶³ REUTER begründet die Beachtung der rechtlichen Aspekte damit, dass die Rechtsbefolgung selbst als moralische Pflicht zu betrachten ist, da das Recht moralische Verantwortungsspielräume gewährt und ein Verhältnis zur moralischen Norm – sei es deckungsgleich, sei es widersprüchlich – inhäriert (vgl. REUTER, *Grundlagen und Methodik der Ethik*, 113). REUTER weiterführend kann in Bezug auf die von ihm angesprochenen Widersprüche auch eine kritische Positionierung gegenüber der rechtlichen Norm selbst als ein Teil der Urteilsfähigkeit betrachtet werden. Wie schon KOHLEBERG in seinen Dilemmageschichten aufzeigt (vgl. Kap. 4.1.2), ist in Bezug auf Divergenzen zwischen der ethischen Pflicht und den Gesetzen die Frage nach der Begründung der Positionierung zu fragen: Wird man rechtsbrüchig mit guter ethischer Begründung oder doch nur aus niedrigen Beweggründen bzw. um seiner Eitelkeit Raum zu geben? In Bezug auf das Rechtswesen ist die Frage fruchtbar, wie ein Staat konstituiert sein müsste, um die Individuen möglichst selten vor eine Wahl zwischen Ethik und Rechtstreue zu stellen.

3. Prüfung, Bewertung und Abwägung

In diesem Arbeitsschritt geht es a) um die Hierarchisierung der handlungsleitenden Orientierungen und b) um die Bewertung der erforderlichen Mittel sowie der zu erwartenden Folgen. Bei konfligierenden Zielen müssen die Vorzugs- und Abwägungsregeln ausgewiesen werden.

4. Entscheidung und Umsetzung

Abschließend sind eine willentliche Zustimmung und praktische Bejahung der Entscheidung (voluntativer Aspekt) notwendig. In einigen Fällen ist zu prüfen, wie mit Spannungen umgegangen werden muss, die aus dem Entscheid resultieren und wie mögliche Kompromisse aussehen könnten, um daraufhin Mittel und Wege der Umsetzung ausfindig zu machen.

Indem die Stufenmodelle definieren, welche Schritte zur Erreichung eines wohl-durchdachten und reflektierten Urteilsentscheides notwendig sind, benennen sie Qualitätsmerkmale, die auch für die Förderung und Bewertung von Urteilskompetenz relevant sind. Diese Art Stufenmodelle werden jedoch in der Fachdidaktik mitunter kritisch diskutiert, da ihnen ein zu stark rationaler und formalistischer Zugang vorgeworfen wird, der Urteilskompetenz auf eine Methodenkompetenz reduziere. Das »Abarbeiten einzelner Schritte« verstelle – so die Kritik – mitunter den inhaltlichen Zugang zur Problemstellung und zum ernsthaften Fragen der Schülerinnen und Schüler und verwehre das Einlassen auf den Inhalt selbst.⁶⁴

Die genannte Kritik hängt jedoch stark von dem didaktischen Umgang mit diesen Modellen ab. Zur Beschreibung des Prozesses liefern sie für Lehrkräfte eine Metastrukturierung, die es ihnen erlaubt, Unterrichtsgespräche zu steuern und Lernprozesse anzuregen. Kritisch hinterfragt werden kann, ob, und wenn ja, in welcher Weise sie den Schülerinnen und Schülern als eine Strukturierung an

⁶⁴ Vgl. in Anlehnung an diese Kritik andere Zugänge, welche sich bewusst dem rationalistischen Paradigma entziehen und auch affektive und emotionale Bestandteile des Urteilens betonen: MIRKA DICKEL, Ethisches Können. Geographieunterricht als Ethisches Können, in: MIRKA DICKEL/ANKE JOHN/MICHAEL MAY/KATHARINA MUTH/LAURENZ VOLKMANN/MARIO ZIEGLER (Hrsg.), Urteilspraxis und Wertmaßstäbe im Unterricht. Ethik, Englisch, Geschichte, politische Bildung und Religion. Schwalbach 2020, 71–99 sowie MARIO ZIEGLER, Die Schulung der Urteilsfähigkeit im Ethikunterricht. Grundsätze der Lehrstückdidaktik, in: MIRKA DICKEL/ANKE JOHN/MICHAEL MAY/KATHARINA MUTH/LAURENZ VOLKMANN/MARIO ZIEGLER (Hrsg.), Urteilspraxis und Wertmaßstäbe im Unterricht. Ethik, Englisch, Geschichte, politische Bildung und Religion. Schwalbach 2020, 47–70.

Vieldiskutiert sind derzeit die Forschungsergebnisse des Psychologen JONATHAN HAIDT, der die primäre Kognitivität von Urteilsentscheiden bezweifelt und demgegenüber die Rolle der Intuitionen und Emotionen in Urteilsbildungsprozessen betont. Er vertritt die Auffassung, dass Urteile emotional getroffen und erst sekundär rational begründet werden (vgl. HAIDT, The Emotional Dog and its Rational Tail).

die Hand gegeben werden sollten. In diesem Fall wäre zumindest ein flexibler Umgang mit diesen Modellen zu betonen und ausreichend Raum für die Fragen der Schülerinnen und Schüler zu lassen, damit der Urteilsprozess ein selbstbestimmter Prozess bleibt. Es darf durch die Modelle nicht zu Einengungen der Schüleransichten und Wertdimensionen kommen, welche MONIKA E. FUCHS vor dem Hintergrund einer im Folgenden noch genauer vorgestellten empirischen Studie resümierend wie folgt beschreibt:

»Datenbasiert kann aufgezeigt werden, in welcher Breite innerhalb einer Lerngruppe gedacht wird und wie unterschiedlich die einzelnen Lernenden Inhalte rezipieren und weiter verarbeiten [sic]. Damit ist zugleich die Herausforderung unterstrichen, vor der Lehrkräfte in der realen Unterrichtssituation stehen. Sie haben jene Palette an Meinungen, Sach- und Wertedimensionen auszuhalten, aufzugreifen und didaktisch fruchtbar zu machen, wobei sich die Struktur subjektiver Urteile und insbesondere deren Normativität auf Seiten Einzelner auch jenseits ethisch intendierter Relevanz bewegen können.«⁶⁵

4.2.3 Empirische Studien zur Argumentations- und Urteilsfähigkeit von Schülerinnen und Schülern

Neben den skizzierten Forschungsbereichen zu Urteilsfähigkeit in der Religionspädagogik gibt es empirische Forschungen, welche vor allem die Argumentationsfähigkeit der Schülerinnen und Schüler in den Blick nehmen. Drei dieser Studien sollen im Folgenden mit Bezug auf die in ihnen benannten Qualitätsmerkmale für Argumentations- und Urteilsfähigkeit genauere Betrachtung finden.

4.2.3.1 Schülerperspektiven auf bioethische Urteilsfragen – Monika E. Fuchs

Eine sehr umfangreiche qualitative Studie mit dem Titel *Bioethische Urteilsbildung im Religionsunterricht* stammt von MONIKA E. FUCHS.⁶⁶ In der konsequent induktiven Erhebung werden Schülerperspektiven auf die Themen »präinatale Diagnostik« und »Behinderung« herausgearbeitet sowie die Auswirkungen von dialogisch-kooperativen bzw. monologisch-individuellen Lehr-Lern-Arrangements auf die Urteilsbildung erforscht.⁶⁷ Auch wenn Qualitäts- und Bewertungskriterien nicht im Zentrum der empirischen Untersuchung FUCHS' stehen, werden durch

⁶⁵ MONIKA E. FUCHS, »Denn irgendwie krank sein bedeutet nicht schlechter sein.« Bioethik aus Schülerperspektive am Beispiel Pränataldiagnostik, in: Zeitschrift für interpretative Schul- und Unterrichtsforschung 1 (2012), 184–199, hier 195 f.

⁶⁶ Vgl. MONIKA E. FUCHS, *Bioethische Urteilsbildung im Religionsunterricht. Theoretische Reflexion – Empirische Rekonstruktion*. Göttingen 2010.

⁶⁷ Vgl. ebd.

die intensive Auseinandersetzung mit Schülerargumentationen Kategorien erhoben, in denen Schülerinnen und Schüler argumentieren. Die Studie bietet folglich in Bezug auf die hier diskutierte Fragestellung nach Qualitätsmerkmalen Hinweise darauf, was von Schülerinnen und Schülern zu erwarten ist. Darüber hinaus sind auch Setzungen der Autorin zu erkennen, beispielsweise wenn Lernarrangements dann als erfolgreich bewertet werden, wenn eine Bewusstseinsweiterung und Sensibilisierung für die Auswirkungen des Urteils auf Seiten der Lernenden stattgefunden haben.⁶⁸ In dieser Setzung ist die bereits genannte Kategorie der Folgenreflexion enthalten.

Im bioethischen Urteilsbildungsprozess ist für FUCHS darüber hinaus der sogenannte *conceptual change* von besonderer Bedeutung: die urteilenden Schülerinnen und Schüler sollen Schritt für Schritt eine wissenschaftliche Sicht auf die Fragestellungen bekommen, während bestehende Alltagsvorstellungen durch wissenschaftlich fundierte Erkenntnisse ergänzt oder auch ersetzt werden (kontinuierliche vs. diskontinuierliche Konzeptwechsel).⁶⁹ Daraus lässt sich ableiten, dass das von FUCHS vertretene Verständnis von Urteilsfähigkeit ein adäquates Sachverständnis über die zu beurteilende Fragestellung als ein Qualitätsmerkmal akzentuiert.

Des Weiteren lassen sich in den Analysen der Schüleraussagen Reflexionen über »Tragweite, Rechtmäßigkeit und Güterabwägung als Bewertungs- und Prozessfaktoren der Urteilsbildung« finden, welche FUCHS als Qualitätsmerkmale der Urteilsbildung deutet.⁷⁰ Hinsichtlich des Urteilsentscheides arbeitet FUCHS Kategorien heraus, anhand derer sich die Schülerinnen und Schüler in ihrem Entscheidungsprozess orientieren: Es werden Handlungsoptionen von Schülerinnen und Schülern abgewogen, denkbare Konsequenzen werden reflektiert, Einflussfaktoren auf den Entscheidungsprozess bedacht und der Begründungshorizont reflektiert, welcher sich zwischen den Polen »individuelle Verortung« und »prinzipielle Verortung« bewegt.⁷¹

4.2.3.2 *Ein Diagnoseinstrument für Argumentationskompetenz in unentscheidbaren, metaphysischen Fragen – Theresa Schwarzkopf*

Eine empirische Studie, welche sich explizit mit der Entwicklung eines lernstandsdiagnostischen Diagnoseinstrumentes für Argumentationsfähigkeit im Bereich der metaphysisch-religiösen Fragen beschäftigt, legte THERESA SCHWARZKOPF

⁶⁸ Vgl. a. a. O., 546.

⁶⁹ Vgl. a. a. O., 245 f.

⁷⁰ Vgl. a. a. O., 596.

⁷¹ Vgl. MONIKA E. FUCHS, Empirische und didaktische Perspektiven bioethischen Lernens im Religionsunterricht, in: Theo-Web. Zeitschrift für Religionspädagogik 8 (2009) 1, 80–90, hier 86.

unter dem Titel *Vielfältigkeit denken. Wie Schülerinnen und Schüler im Religionsunterricht argumentieren lernen* vor (zur Lernstandsdiagnostik vgl. Kap. 3.3.2).⁷² Die Studie untersucht die Lernstufen der Argumentationsfähigkeit in Bezug auf die Auferstehung der Toten und Gottesvorstellungen. Mit diesem Fokus auf »unentscheidbare«, metaphysische Fragen hat SCHWARZKOPF einen Gegenstand gewählt, der sich Sachargumentationen, aber auch anderen gängigen Urteilskriterien wie der Folgenreflexion entzieht (vgl. Kapitel 2.1). Dennoch hebt SCHWARZKOPF die Bewertbarkeit dieser – stark von individuellen Glaubensüberzeugungen abhängigen – Urteilsfindungen hervor, indem sie schreibt, dass »die Begründung der gewählten Antwortoptionen in ihrer inneren Logik und ihrer argumentativen Abgrenzung zu den weiteren Antwortoptionen sehr wohl bewertet werden und somit effektives Lernen begründen« kann.⁷³

SCHWARZKOPF geht davon aus, dass die eigenen Konstrukte untrennbar mit Formulierungs- und Argumentationskompetenzen verknüpft sind und sich gegenseitig bedingen.⁷⁴ Als Qualitätsmerkmale der Argumentation bestimmt sie zum einen die Generierung einer eigenen Position und zum anderen die nachvollziehbare Darlegung dieser.⁷⁵ Sie erarbeitet dazu folgende Komponenten: Inhaltliche Kohärenz, Argumente auf Modellebene (d. h. die Durchdringung fremder Modelle und die distanzierte Betrachtung der eigenen Argumente als Modelle), die Möglichkeit der Perspektivübernahme anderer Positionen sowie die sachbezogene Angemessenheit der Argumente und Schlussfolgerungen.⁷⁶ Das von ihr zur Erfassung der Argumentationsfähigkeit der Schülerinnen und Schüler entwickelte Diagnoseverfahren bildet die Bearbeitungsstruktur ab. Urteilsbildung wird dabei als Ziel der Argumentation von SCHWARZKOPF stets mitgedacht.

⁷² Vgl. SCHWARZKOPF, *Vielfältigkeit denken*.

⁷³ Vgl. a. a. O., 54.

⁷⁴ Vgl. a. a. O., 25 f.

⁷⁵ Vgl. a. a. O., 86.

⁷⁶ Vgl. a. a. O., 178.

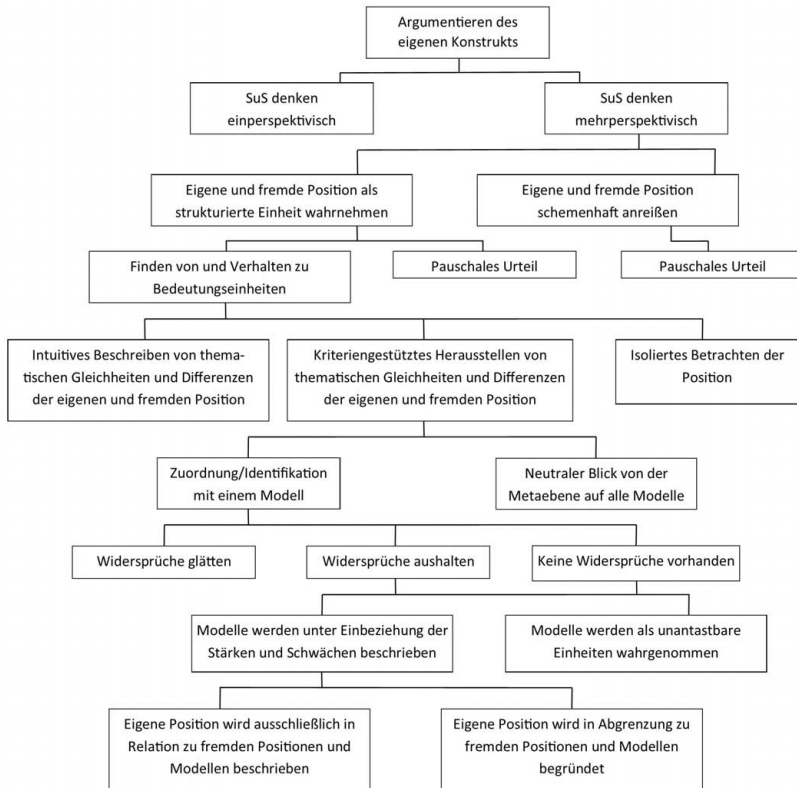


Abb. 9: Bearbeitungsstruktur zur Argumentationskompetenz in metaphysisch-religiösen Fragen. Nachbildung aus SCHWARZKOPF: Vielfältigkeit denken, 179.

Als empirisch abbildbare, normative Lernschritte im Prozess des Argumentieren-Lernens benennt SCHWARZKOPF 1.) die mehrperspektivische Denkweise, 2.) die strukturierte Wahrnehmung fremder und eigener Positionen, 3.) affektives und differenziertes Beurteilen, 4.) eine kriteriengestützte Bezugnahme und 5.) die Argumentation auf Modellebene.⁷⁷ Die erhobenen Lernschritte überträgt SCHWARZKOPF in ein Niveaustufenmodell, um neben einer Lernprozessdiagnostik auch eine Lernstandsdiagnostik zu ermöglichen. Die höchste der fünf Stufen definiert SCHWARZKOPF folgendermaßen:

»Die SuS [Schülerinnen und Schüler] argumentieren für ihre Vorstellung, indem sie diese und die fremden Positionen systematisch in ihren einzelnen Bedeutungseinheiten vorstellen. Die eigene und fremde Vorstellung werden [sic] in einem religiösen

⁷⁷ Vgl. a. a. O., 183.

Denkmodell rekonstruiert, wobei hier auftretende Widersprüche ausgehalten werden. Auf Grundlage der Modellrekonstruktion analysieren die SuS Gemeinsamkeiten, Differenzen und Widersprüche zwischen den Positionen bzw. Modellen und zeigen so metastrukturelle Beziehungen sowie Stärken und Schwächen der Modelle auf. Aufbauend auf dieser Analyse begründen die SuS ihre Vorstellung in Abgrenzung zu den fremden Positionen und Modellen, um diese rational nachvollziehbar zu machen.«⁷⁸

Zusammenfassend können nach SCHWARZKOPF in Bezug auf metaphysische Fragestellungen die Anerkennung von Mehrperspektivität im Sinne von Gleichberechtigung verschiedener Perspektiven, eine strukturierte Wahrnehmung fremder und eigener Positionen sowie das Aushalten von Widersprüchen zu den fremden Positionen als Qualitätsmerkmale des Urteilsbildungsprozesses benannt werden.

4.2.3.3 *Argumentationstheorien in der Praxis – Thomas Weiß*

Eine Studie, die sich ebenso mit der Argumentationsfähigkeit auseinandersetzt, stammt von THOMAS WEIß. Er hat exemplarisch am Thema Schöpfung und Evolution empirisch erhoben, welches fachspezifische und fachübergreifende Argumentationspotential 16- bis 18-jährige Schülerinnen und Schüler im Kontext der Schulfächer Religion und Biologie besitzen.⁷⁹ Ausgehend von der Typologie kontextabstrakter Argumentationsmuster nach MANFRED KIENPOINTNER, welche im Gegensatz zu anderen Argumentationstypologien an natürlichen Dialogsituationen ansetzt und daher auch geeignet ist, alltagsnahe Schüleraussagen zu fassen, entwickelt WEIß ein Diagnoseinstrument zur Klassifikation verwendeter Schlussregeln.⁸⁰ Den Weg vom Grund zur Konklusion klassifiziert KIENPOINTNER folglich »nicht nach der Komplexität der Argumentation, d. h., er fragt nicht, aus wie vielen Argumenten eine Argumentation besteht, sondern nach der Art der Verwendung der Schlussregel.«⁸¹

WEIß kommt zu dem Ergebnis, dass in den Texten gymnasialer Oberstufenschülerinnen und -schüler zwar argumentiert wird und dass sich in diesen auch die von KIENPOINTNER systematisierten Argumentationsmuster aller drei Großklassen der Typisierung nachweisen lassen, dass es jedoch für die untersuchten Schülerinnen und Schüler eine Schwierigkeit zu sein scheint,

⁷⁸ Vgl. a. a. O., 188.

⁷⁹ Vgl. THOMAS WEIß, Fachspezifische und fachübergreifende Argumentationen am Beispiel von Schöpfung und Evolution. Theoretische Grundlagen – empirische Analysen – Jugendtheologische Konsequenzen. Göttingen 2016.

⁸⁰ Vgl. KIENPOINTNER, Alltagslogik.

⁸¹ A. a. O., 42.

»eine in sich stimmige, die einzelnen Haltungen und Überzeugungen abbildende Argumentation vorzutragen. Es kann zwar notiert werden, dass die Schülerinnen und Schüler im Durchschnitt die Anstrengung unternehmen, Fragen aufzuwerfen, Perspektivenwechsel zu vollziehen und Lösungsvorschläge zu unterbreiten. Dies darf aber nicht darüber hinwegtäuschen, dass eine begriffliche Inkonsistenz besteht, die sich letztlich in den meisten hier [im Rahmen der Studie von Weiß; Anm. KM] besprochenen Argumentationen niederschlägt.«⁸²

Darüber hinaus stellt WEIß fest, dass – entgegen seiner angenommenen Hypothese, dass im theologischen Bereich die Verwendung normativer Argumentationen höher sei als im naturwissenschaftlichen – die Schülerinnen und Schüler fachunabhängig überwiegend deskriptiv anstatt normativ argumentieren.⁸³

4.3 Zusammenfassung

Der Einblick in den aktuellen Forschungsstand macht deutlich, dass Argumentations- und Urteilsfähigkeit zu einem Forschungsbereich gehören, der unter verschiedenen Perspektiven beforscht wird und, dass vor allem mit Blick auf didaktische Prozesse, Niveaustufen entwickelt wurden, um die Qualität von Urteilsbildungsprozessen messbar zu machen. Dabei wird seit KOHLBERGS Studie über die Moralentwicklung nicht mehr in Frage gestellt, dass die Positionierung zu einer Option nichts über die Qualität des Urteils aussagt, sondern dass sich diese über andere Merkmale generiert. Folgende Tabelle soll die Qualitätsmerkmale, die in den vorgestellten empirischen Forschungen herausgearbeitet wurden, in einem Überblick zusammenstellen.

⁸² Vgl. a.a.O., 474.

⁸³ Vgl. a.a.O., 477.

Dieser Befund deckt sich mit Ergebnissen aus der Geschichtsdidaktik: SCHÖNEMANN, THÜNEMANN und ZÜLSDORF-KERSTING weisen in einer Untersuchung von Abituraufgabensätzen nach, dass normative Werturteile – im Gegensatz zu Sachurteilen – von den Schülerinnen und Schülern kaum gefällt werden (vgl. DIES., Was können Abiturienten?, 68).

<i>Vertreter/in</i>	<i>Modell/Forschung</i>	<i>Qualitätsmerkmale ethischer Urteilskompetenz</i>
LAWRENCE KOHLBERG	Stufen der moralischen Entwicklung	– Ausrichtung an der Gesellschaft vorgeordneten (nützlichen oder universal ethischen) Prinzipien
GEORG LIND	»Moralische-Kompetenz-Test« (MKT) sowie Konstanzer Methode der Dilemmadiskussion	– Prüfung der moralischen Qualität (aufbauend auf KOHLBERG) von Pro- und Kontra-Argumenten entgegengesetzt der Meinungskonformität (Multiperspektivität)
DIETRICH BENNER, ROUMIANA NIKOLOVA <i>et al.</i>	Berliner Ansatz zur Konstruktion und Erhebung ethisch-moralischer Kompetenzniveaus im öffentlichen Erziehungs- und Bildungssystem (ETiK)	– Urteilen auch in lebensweltlich nicht vertrauten Problemlagen – Reflexion und Prüfung gesellschaftlicher Werte
ELKE VISSER	Diagnose der Bewertungskompetenz durch schriftliche Aufgaben im Biologieunterricht	– Berücksichtigung der Perspektiven auch mittelbar Beteiligter – Folgenreflexion unter Berücksichtigung der gesellschaftlichen Ebene – Transparenz der dem Urteil zugrunde liegenden Wertungen, Normen und Werte – Abwägen von Argumenten unterschiedlicher Positionen (Multiperspektivität)
MONIKA E. FUCHS	Bioethische Urteilsbildung im Religionsunterricht	– Beachtung der Auswirkungen des Urteils – adäquates Sachverständnis – Tragweite (Folgenreflexion), Rechtmäßigkeit und Güterabwägung als Reflexionskriterien der Urteilsbildung

Fortsetzung

<i>Vertreter/in</i>	<i>Modell/Forschung</i>	<i>Qualitätsmerkmale ethischer Urteilskompetenz</i>
THERESA SCHWARZKOPF	metaphysisch-religiöse Fragen Vielfältigkeit denken. Wie Schülerinnen und Schüler im Religionsunterricht argumentieren lernen	- Anerkennung von Mehrperspektivität im Sinne von Gleichberechtigung verschiedener Perspektiven - strukturierte Wahrnehmung fremder und eigener Positionen - Aushalten von Widersprüchen zu den fremden Positionen

Tab. 7: Empirisch überprüfte Qualitätsmerkmale ethischer und religiöser Urteilskompetenz im Überblick.

Aus der Übersicht wird deutlich, dass Mehrperspektivität, also das Wahrnehmen und Reflektieren gegensätzlicher Positionen, ein wesentlicher Stellenwert in Hinblick auf die Frage eingeräumt wird, was ein reflektiertes Urteil ausmacht.

Vergleicht man die von SCHWARZKOPF herausgearbeiteten Qualitätsmerkmale für Urteilsbildung in metaphysisch-religiösen Fragen mit den theoretischen und empirischen Forschungen zu ethischer und religiöser Urteilsbildung, wird deutlich, dass sich unterschiedliche Gewichtungen benennen lassen. Einigen Qualitätsmerkmalen ethisch-philosophischer und ethisch-religiöser Urteilsbildung kommen im Bereich metaphysisch-religiöser Fragestellungen nur eine sehr marginale bis gar keine Bedeutung zu wie der Folgenabwägung, der Prüfung der rechtlichen Rahmen und der Umsetzbarkeit oder der Kenntnisaufnahme empirischer Fakten. Ebenso tritt die KOHLBERGS Modell zugrunde liegende Ausrichtung an der Gesellschaft vorgeordneten (nützlichen oder universal ethischen) Prinzipien oder die von VISSER herausgearbeitete Transparenz der zugrunde liegenden Normen und Werte in Bezug auf metaphysisch-religiöse Fragen in den Hintergrund, da – wie eingangs in der Systematisierung dargestellt (Kap. 2.1) – Normen und Werte mitunter aus einer spezifischen Weltsicht und Sinndeutung resultieren können, sich metaphysische Konstrukte jedoch nicht zwingend aus gesellschaftlichen Normen und Werten ableiten lassen müssen, sondern oft auf religiösen Texten, Traditionen und Sozialisationen gründen. Sie haben dadurch in Bezug auf die rechtfertigende Argumentation und den Allgemeingültigkeitsanspruch einen anderen Status. Wie SCHWARZKOPF zeigt, treten andere Qualitätskriterien, nämlich die innere Kohärenz eines Konstrukts, die Anerkennung verschiedener Perspektiven als gleichberechtigt sowie das Aushalten von Widersprüchen in den Vordergrund. Argumentationsfähigkeit und die Auseinandersetzung mit anderen Positionen sind hingegen in beiden Bereichen gleichermaßen gefordert.

In der Darstellung des aktuellen Forschungsstandes zur ethischen und religiösen Urteilsfähigkeit wurden Qualitätsmerkmale für Urteilsbildung herausgearbeitet, welche zum Teil implizit, zum Teil aber auch explizit als solche im Forschungsdiskurs angenommen und diskutiert werden. Dadurch konnte das in der Forschungsliteratur vorhandene Konstrukt von Urteilsfähigkeit näher beschrieben werden. Diese Erkenntnisse aus der Forschung sollen als Folie verwendet werden, um die in den schriftlichen Abituraufstellungen und Erwartungshorizonten benannten Qualitätsmerkmale in den Forschungsdiskurs einordnen zu können. Im Folgenden wird zunächst das Forschungsdesign der Erhebung beschrieben.

5 Forschungsdesign

Das Kapitel Forschungsdesign dient der Transparenz der verwendeten Auswahlkriterien für die der Studie zugrunde liegenden Materialien (Kapitel 5.1). Darüber hinaus wird das methodische Vorgehen erläutert und begründet, die Auswahl der Analyseeinheiten definiert und im Material zu findende Grenz- und Sonderfälle diskutiert (Kapitel 5.2). Daraufhin findet eine Vorstellung des entwickelten Kategoriensystems statt, welches für die Inhaltsanalyse der Abituraufgabenstellungen und ihrer Erwartungshorizonte deduktiv entwickelt und im Zuge der Analyse induktiv ergänzt wurde (5.3). Abschließend wird das Analyseverfahren beschrieben und die Wahl der Methodik reflektiert (5.4). Zeitlich betrachtet fand der Prozess der Methodenreflexion während und nach der Inhaltsanalyse und vor allem vor dem Hintergrund ihrer Ergebnisse statt. Die Zuordnung der Methodenreflexion in das Kapitel Forschungsdesign folgt somit keiner chronologischen, sondern einer inhaltlichen Gliederung.

5.1 Material

Als Untersuchungsmaterial zur Beantwortung der Fragestellung, wie Urteilsbildung in ethischen und religiösen Fragen geprüft und welche bewertungsrelevanten Qualitätsmerkmale in Erwartungshorizonten genannt werden, dienen die im Zeitraum von 2014 bis 2019 gestellten schriftlichen Abituraufgabenstellungen aus den exemplarisch ausgewählten Bundesländern Bayern, Thüringen und Niedersachsen im Fach Evangelische Religionslehre sowie die dazugehörigen Erwartungshorizonte. Mit diesem Material rücken die ministerial verantworteten Prüfungskonventionen in den Blick. Das Material gibt keine Auskunft über die Aufnahme der Aufgaben seitens der Prüflinge oder über die Umsetzung der Bewertungsvorgaben. Obwohl der Personenkreis derer, die die Abituraufgabenstellungen und die Erwartungshorizonte erstellen, begrenzt ist und davon auszugehen ist, dass das Material auch von individuellen Annahmen in Bezug auf Bildung und Urteilsbildung geprägt sein kann, hat es doch vor dem

Hintergrund des Zentralabiturs und der angestrebten Vergleichbarkeit von Bildungsabschlüssen erhebliche Weisungskraft und normierende Wirkung hinsichtlich der Qualitätsbestimmungen von Lernerfolgen (zur Erstellung der Aufgaben in Fachkommissionen vgl. Kap. 3.1).

Durch die exemplarische Auswahl dreier Bundesländer soll darüber hinaus die Vergleichbarkeit zwischen den länderspezifischen Zentralabituren untersucht werden. Die Auswahl der Bundesländer folgt dabei mehreren Differenzkriterien: Zum einen führt Thüringen hinsichtlich der Gesamtnotendurchschnittswerte der Abschlusszeugnisse über Jahre hinweg die Spitze der Bundesländer an, während Niedersachsen – ebenso konstant – im Mittel die schlechtesten Abiturnoten verteilt.¹ Bayern liegt zwischen den beiden Bundesländern, wie folgende Tabelle zeigt:

	Thüringen	Bayern	Niedersachsen
2019	2,18	2,32	2,56
2018	2,16	2,31	2,57
2017	2,18	2,31	2,57
2016	2,18	2,32	2,58
2015	2,16	2,31	2,59
2014	2,16	2,33	2,61

Tab. 8: Gesamtdurchschnitt der Abiturnoten in den Bundesländern Bayern, Niedersachsen und Thüringen. Eigene Tabelle auf Grundlage des Sekretariats der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland IVC/Statistik.

Unterschiede in den Durchschnittsnoten wären an sich nicht ungewöhnlich, da die Leistungen der Schülerinnen und Schüler durchaus schwanken könnten. Jedoch bleiben die Verteilungen über mehrere Jahre gleich, was darauf hinweist, dass andere Faktoren als die Leistung der einzelnen Schülerinnen und Schüler den Ausschlag geben. Dies ist vor allem vor dem Hintergrund bundesweit gültiger Hochschulzugangsbedingungen problematisch (vgl. Kap. 3.1). Die Gründe für die starken und über Jahre hinweg gleichbleibenden Differenzen der Abschlussnoten können vielfältig sein und dürfen nicht eindimensional zum Beispiel auf die Qualität des erteilten Unterrichts oder auf die Intelligenz der

¹ Vgl. STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BUNDESREPUBLIK DEUTSCHLAND IVC/STATISTIK (Hrsg.), Abiturnoten 2019 an Gymnasien, Integrierten Gesamtschulen, Fachgymnasien, Fachoberschulen und Berufsschulen. Abrufbar unter: <https://www.kmk.org/dokumentation-statistik/statistik/schulstatistik/abiturnoten.html> (07.2.2020). Wie bereits aus dem Titel hervorgeht, schließt der Vergleich neben »den Absolventen der allgemeinbildenden Schularten Gymnasium und integrierte Gesamtschule auch die Absolventen der Fachgymnasien, Fachoberschulen und Berufsoberschulen« mit ein (vgl. a.a.O.).

Schülerinnen und Schüler zurückgeführt werden. So unterscheiden sich beispielsweise die Anzahl der Aufgabenstellungen und die Bearbeitungszeiten der Abiturprüfungen zwischen den Bundesländern (vgl. Kapitel 6.1.1). Darüber hinaus gibt es Unterschiede in den Oberstufensystemen hinsichtlich der Gewichtung einzelner Fächer für die Gesamtnote und – da die Statistik mehrere Schulformen zusammenfasst – können Unterschiede auch durch die Zusammensetzung der Schullandschaft bedingt sein.²

Ein weiteres Differenzkriterium zur Auswahl der Bundesländer besteht hinsichtlich der religiösen Sozialisierung der Bevölkerung und insbesondere der Schülerschaft, die sich zwischen West- und Ostdeutschland deutlich unterscheidet. Der Religionsmonitor der Bertelsmann-Stiftung spricht von einer »bestehende[n] religiöse[n] Kluft zwischen West- und Ostdeutschland«³:

»Glaubt im Westen etwa jeder Zweite »ziemlich« beziehungsweise »sehr« daran, dass Gott, Gottheiten oder etwas Gottähnliches existiert, tut dies im Osten nur knapp jeder Vierte. Während der Anteil der eher Gläubigen im Westen doppelt so hoch ist wie der der eher nicht Gläubigen [...], stellen Letztere im Osten mit knapp 70 % die übergroße Mehrheit.«⁴

Mit Thüringen ist den Bundesländern Bayern und Niedersachsen ein stark säkularisiertes Bundesland an die Seite gestellt: Ca. 70 % der Bevölkerung hat »(mit regionalen Schwankungen) teilweise bereits in der 4. Generation keinerlei Bezüge zu einer Religionsgemeinschaft [...] oder [empfindet] sich noch als religiös.«⁵ Dies wirkt sich auch auf die Schule aus. MICHAEL WERMKE und MARIA WIDL schreiben in Bezug auf religiöse Angebote im Schulleben:

»Das Thüringer Kultusministerium verfolgt seit vielen Jahren – unabhängig von der Parteizugehörigkeit der jeweiligen Kultusministerinnen und -minister – einen relativ strikten Kurs der Trennung von Kirche und Schule, die die an den Schulen nach wie vor verbreitete Skepsis gegenüber religiösen Angeboten im Schulleben eher stärkt.«⁶

² Gymnasien, Integrierte Gesamtschulen, Fachgymnasien, Fachoberschulen und Berufsschulen.

³ DETLEF POLLACK/OLAF MÜLLER, Religionsmonitor. Verstehen was verbindet. Religiosität und Zusammenhalt in Deutschland. Gütersloh 2013, 10.

⁴ A. a. O., 10 f.

⁵ MICHAEL WERMKE/MARIA WIDL, Religion unterrichten in Thüringen, in: MARTIN ROTHGANGEL, BERND SCHRÖDER (Hrsg.), Religionsunterricht in den Ländern der Bundesrepublik Deutschland. Neue empirische Daten – Kontexte – Aktuelle Entwicklungen. Leipzig 2020, 419–442, hier 437.

⁶ A. a. O., 432.

Von den Schülerinnen und Schülern, die den evangelischen Religionsunterricht besuchen, ist ein beachtlicher Anteil – nämlich 25 % – konfessionell nicht gebunden.⁷

Dass das Fach Religion in den Bundesländern auf unterschiedliche Akzeptanz in Schule und Gesellschaft stößt, wird auch an den Belegungszahlen sichtbar. In Bayern besuchten im Schuljahr 2017/18 insgesamt ~45 % den katholischen und evangelischen Religionsunterricht.⁸ In Niedersachsen waren es im gleichen Schuljahr ~32 %, während es in Thüringen nur 18 % waren.⁹

<i>Schuljahr 2017/18</i>	<i>Teilnehmer*innen ev. RU in allen Schularten¹⁰</i>	<i>Teilnehmer*innen kath. RU in allen Schularten¹¹</i>	<i>Gesamtschüler* innenzahl¹²</i>
Bayern	~14 % (236.659)	~31 % (523.051)	1.672.312
Niedersachsen	~27 % (304.161)	~5 % (58.561)	1.122.378
Thüringen	~14 % (34.014)	~3 % (9.109)	242.042

Tab. 9: Vergleich der Teilnehmerinnen und Teilnehmerzahlen am Religionsunterricht in Bayern, Niedersachsen und Thüringen im Schuljahr 2017/18.

Auch wenn die drei Bundesländer nicht repräsentativ für die gesamte Bundesrepublik stehen können, bilden sie doch drei in Deutschland wesentliche und differente religiöse Milieus ab: Bayern ist überwiegend katholisch geprägt, Niedersachsen in erster Linie protestantisch und mit Thüringen kommt ein überwiegend säkular geprägtes Umfeld in den Blick.

Auf eine unterschiedliche Gewichtung des Schulfaches Religion in den drei Bundesländern weisen auch die Differenzen bezüglich der Abnahme schriftlicher

⁷ Vgl. ebd.

⁸ Vgl. STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BUNDESREPUBLIK DEUTSCHLAND IVC/STATISTIK (Hrsg.), Auswertung Religionsunterricht Schuljahr 2017/18. Teilnehmende Schülerinnen und Schüler allgemeinbildender Schulen in öffentlicher Trägerschaft nach Schularten (aufgegliedert nach Religionsunterricht, Ethik und weiteren Ersatzunterrichten) für den Primar- und Sekundarbereich I. Berlin 2019. Abrufbar unter: https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/AW_Religionsunterricht_II_2017_18.pdf (26.3.2019).

⁹ Die Zahlen aus dem Schuljahr 2017/2018 sind derzeit – im Frühjahr 2019 – die aktuellsten verfügbaren Zahlen.

¹⁰ Vgl. ebd.

¹¹ Vgl. ebd.

¹² STATISTISCHES BUNDESAMT (Hrsg.), Pressemitteilung. Schülerzahlen im Schuljahr 2017/18 um 0,4 % zurückgegangen. Abrufbar unter: https://www.destatis.de/DE/Presse/Pressemitteilungen/2018/03/PD18_089_211.html (26.3.2020).

Abiturprüfung im Fach Evangelische Religionslehre hin. Da in Thüringen ausschließlich schriftliche Abiturprüfungen in Fächern mit erhöhtem Anforderungsniveau abgelegt werden können und dieses erhöhte Anforderungsniveau im Fach Evangelische Religionslehre nur an einer Schule in ganz Thüringen angeboten wird (mit dem Fach Ethik verhält es sich adäquat), sind die Zahlen der Schülerinnen und Schüler, die im Fach Evangelische Religionslehre eine schriftliche Prüfung ablegen, sehr gering, wie folgende Tabelle verdeutlicht.¹³

<i>Schuljahr</i>	<i>Anzahl der Prüflinge Thüringen</i>	
	<i>Evangelische Religionslehre</i>	<i>Ethik</i>
2013/2014	8	3
2014/2015	4	11
2015/2016	10	4
2016/2017	19	10
2017/2018	8	22
2018/2019	17	15

Tab. 10: Anzahl der Prüflinge schriftlicher Abiturprüfungen in den Fächern Evangelische Religionslehre und Ethik im Bundesland Thüringen.

Die Gestaltung der Aufgabenstellungen und ihrer Erwartungshorizonte sollten von der geringen Anzahl der Prüflinge unbeeinflusst bleiben, da es sich – ebenso wie bei den anderen beiden Bundesländern – um ministerial verantwortete zentrale Abiturprüfungen handelt, die auch bei einer geringen Adressatenzahl den gleichen qualitativen Ansprüchen gerecht werden sollten, wie in Bundesländern mit einer großen Adressatenzahl. Die geringe Zahl an Prüflingen wurde daher nicht als Grund gesehen, Thüringen aus der Untersuchung herauszunehmen. Gegebenenfalls kann der Adressatenkreis der Prüfung von der Prüfungskommission besser eingeschätzt werden als in Bundesländern, in denen die Prüfungen für viele Schulen erstellt werden. Dennoch lassen sich diese Einschätzungen nicht generalisieren, so dass Objektivität, Validität und Reliabilität

¹³ Die Daten wurden dankenswerter Weise auf Anfrage vom THÜRINGER MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT zusammengestellt. Zum Thüringer Oberstufensystem vgl. MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT (Hrsg.), *Gymnasiale Oberstufe in Thüringen. Voraussetzungen, Struktur, Unterrichtsfächer, Bewertung und Abiturprüfung*. https://bildung.thueringen.de/fileadmin/ministerium/publikationen/gymnasiale_Oberstufe.pdf (13.5.2020).

Die Differenzierung zwischen erhöhtem und grundlegendem Anforderungsniveau wurde in einem Beschluss der Kultusministerkonferenz vom 7.7.1972 in der Fassung vom 2.6.2006 festgelegt und wird auch durch die EPA vorgeschrieben (vgl. EPA, 11). Die Umsetzung und die Prüfungsmöglichkeiten sind jedoch bundeslandspezifisch.

auch für einen kleinen Adressatenkreis anzustrebende Kriterien sind (vgl. Kapitel 3.3.2). So wäre – gerade auch für den Fortbestand des Religionsunterrichts – bedeutsam, dass sich der gesellschaftliche Bedeutungsverlust des Faches (siehe oben) nicht negativ auf die Qualität des Faches auswirkt.¹⁴

In Bayern hingegen zeichnet sich in Bezug auf schriftliche Abiturprüfungen im Fach Evangelische Religionslehre ein anderes Bild. Dort sind die Möglichkeiten und damit auch die Anzahl an Prüflingen, die eine Prüfung ablegen, weitaus höher, wie folgende Tabelle zeigt.¹⁵ Bayern differenziert im Fach Evangelische Religion nicht zwischen erhöhtem und grundlegendem Anforderungsniveau.

<i>Schuljahr</i>	<i>Anzahl der Prüflinge Bayern</i>	
	<i>Evangelische Religionslehre</i>	<i>Ethik</i>
2013/2014	385	371
2014/2015	339	357
2015/2016	342	347
2016/2017	320	304
2017/2018	330	314
2018/2019	339	281

Tab. 11: Anzahl der Prüflinge schriftlicher Abiturprüfungen in den Fächern Evangelische Religionslehre und Ethik im Bundesland Bayern.

In Niedersachsen ist das Ablegen einer schriftlichen Abiturprüfung im Fach Evangelische Religionslehre in zwei verschiedenen Anforderungsniveaus möglich: Der Unterricht auf erhöhtem Anforderungsniveau umfasst fünf Wochenstunden und ist mit einem gesellschaftswissenschaftlichen Schwerpunkt verknüpft.¹⁶ Er soll »in besonderem Maße der allgemeinen Studienvorbereitung« dienen und »in wissenschaftliche Methoden, Fragestellungen und Reflexionen

¹⁴ Dies erscheint – neben aller pädagogischen und bildungstheoretischen Notwendigkeit – vor allem vor dem Hintergrund sinkender Studierendenzahlen in Theologie und Lehramt evangelische Religion zentral zu sein, trägt doch die Schule zu dem Verständnis der Fachdisziplinen bei Schulabsolventinnen und -absolventen wesentlich bei.

¹⁵ Die Daten wurden auf Anfrage vom BAYERISCHES STAATSMINISTERIUM FÜR UNTERRICHT UND KULTUS zusammengestellt. Auch hierfür ist ausdrücklich zu danken.

¹⁶ Vgl. NIEDERSÄCHSISCHES KULTUSMINISTERIUM (Hrsg.), Die gymnasiale Oberstufe und die Abiturprüfung. Informationen für Eltern und die Schülerinnen und Schüler, die ab 2021 ihr Abitur ablegen werden. Hannover: 2019, 5f. Das den Angaben zugrunde liegende Schulgesetz ist am 01.08.2015 in Kraft getreten, so dass sich die Informationen weitestgehend auf den Untersuchungszeitraum beziehen lassen.

eingeführen¹⁷. Religionsunterricht auf grundlegendem Anforderungsniveau wird in drei Wochenstunden erteilt.¹⁸ Er kann unabhängig von der Schwerpunktsetzung besucht werden und verfolgt die Vermittlung grundlegender Sachverhalte, Fähigkeiten und Fertigkeiten. Das Pendant zum Religionsunterricht nennt sich Werte und Normen und kann schriftlich nur auf grundlegendem Anforderungsniveau geprüft werden. Über die Inanspruchnahme der Möglichkeit, sich im Fach Evangelische Religionslehre und im Fach Werte und Normen prüfen zu lassen, gibt folgende Tabelle Auskunft:

<i>Schuljahr</i>	<i>Anzahl der Prüflinge Niedersachsen</i>		
	Evangelische Religionslehre (grundlegendes Anforderungsniveau)	Evangelische Religionslehre (erhöhtes Anforderungsniveau)	Werte und Normen (grundlegendes Anforderungsniveau)
2013/2014	389	458	62
2014/2015	412	587	66
2015/2016	371	592	55
2016/2017	419	488	118
2017/2018	375	493	131
2018/2019	389	549	135

Tab. 12: Anzahl der Prüflinge schriftlicher Abiturprüfungen in den Fächern Evangelische Religionslehre und Werte und Normen im Bundesland Niedersachsen.

Bei den in den Tabellen angeführten Zahlen handelt es sich um absolute Zahlen, welche mit der Gesamtzahl der Abiturientinnen und Abiturienten eines Jahrgangs in Bezug gesetzt werden müssen, da Thüringen hinsichtlich der Fläche und der Bevölkerungszahl zu den kleinsten Bundesländern gehört, während Bayern und Niedersachsen zu den größten Bundesländern zählen. Dennoch sind die Differenzen in Bezug auf die Möglichkeit und Wahrnehmung, ein schriftliches Abitur im Fach Evangelische Religionslehre abzulegen, so gravierend, dass sie einen Hinweis auf eine unterschiedliche gesellschaftliche Anerkennung des Faches in den drei Bundesländern geben.

5.2 Methode

Die Analyse der Aufgabenstellungen und der dazugehörigen Erwartungshorizonte erfolgt methodisch in zwei Schritten. Der erste Schritt besteht aus einer

¹⁷ NIEDERSÄCHSISCHES KULTUSMINISTERIUM (Hrsg.), Die gymnasiale Oberstufe, 5.

¹⁸ Vgl. ebd.

qualitativen Inhaltsanalyse wie sie von PHILIPP MAYRING und von UDO KUCKARTZ im Repertoire sozialwissenschaftlicher Methoden etabliert wurde.¹⁹ Im Vordergrund dieser Methode steht ein systematisches Vorgehen anhand expliziter Analyseregeln mit dem Ziel, die Nachvollziehbarkeit und Überprüfbarkeit der Analyse zu gewährleisten.²⁰ Das zentrale Instrument stellt dabei ein Kategoriensystem dar, welches am konkreten Gegenstand entwickelt wird und die Regeln der Analyse transparent macht.²¹ Das Kategoriensystem der vorliegenden Studie basiert auf der inhaltlich strukturierenden qualitativen Inhaltsanalyse nach UDO KUCKARTZ, welche in erster Linie der »Identifizierung von Themen und Subthemen, deren Systematisierung und Analyse der wechselseitigen Relationen« dient.²² Ziel ist eine Analyse und Strukturierung der aktuell gängigen Prüfungskonventionen und Bewertungsinstruktionen von Urteilsfähigkeit. Somit wurde sich bewusst gegen eine evaluative qualitative Inhaltsanalyse entschieden, bei der normative Bewertungen bereits im Kategoriensystem gesetzt und an das Material herangetragen werden, um dieses einzuschätzen.²³ Bei der inhaltlich strukturierenden qualitativen Inhaltsanalyse wird das Material durch das Kategoriensystem zunächst systematisiert. Eine Bewertung des Materials findet erst im Kontext der Ergebnisdiskussion statt.

Die Hauptkategorien sind deduktiv, also unabhängig von der empirischen Datenbasis, entstanden.²⁴ Sie leiten sich in erster Linie vom Forschungsanliegen

¹⁹ Vgl. PHILIPP MAYRING, *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim/Basel ¹²2015 sowie UDO KUCKARTZ, *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Weinheim/Basel ⁴2018.

²⁰ Vgl. MAYRING, *Qualitative Inhaltsanalyse*, 12 f. Darüber hinaus sind die theoretische Herleitung der Fragestellung sowie die Interpretation der Ergebnisse vor dem jeweiligen Theoriehintergrund relevant, um Schlussfolgerungen für den Kommunikationskontext ziehen zu können, aus welchem das Material stammt (vgl. a. a. O., 13). Die in der vorliegenden Studie untersuchten Materialien stehen in einem Kommunikationsfeld zwischen Aufgabensteller bzw. Aufgabenstellerin, Prüfling und korrigierenden Lehrkräften. Die untersuchten Materialien müssen bestenfalls in der Lage sein, die Kommunikation in diesem Kontext passgenau zu steuern.

²¹ Vgl. MAYRING, *Qualitative Inhaltsanalyse*, 51 f.

²² Vgl. KUCKARTZ, *Qualitative Inhaltsanalyse*, 123.

Zur inhaltlich strukturierenden Qualitativen Inhaltsanalyse vgl. a. a. O., 97–121.

²³ Vgl. a. a. O., 123–141.

²⁴ KUCKARTZ spricht in Zusammenhang mit der »deduktiven Kategorienbildung« auch von einer »A-priori-Kategorienbildung«. Charakteristisch für diese Art von Kategorienbildung ist, dass sie vor der Sichtung des Datenmaterials stattfindet also theoriegeleitet aus dem Forschungsstand abgeleitet wird (vgl. UDO KUCKARTZ, *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Weinheim/Basel ⁴2018, 64). »Die Kategorien fungieren als eine Art Suchraster, d. h. das Material wird auf das Vorkommen des entsprechenden Inhalts durchsucht und grob kategorisiert.« (vgl. a. a. O., 96).

ab (vgl. Kap. 2.4) und berücksichtigen a) den in Kapitel 4 dargestellten Forschungsstand zur Urteilsbildung, b) die bildungsadministrativen Vorgaben der Einheitlichen Prüfungsanforderungen (EPA; vgl. Kap. 3.2) sowie c) die im Bereich der Aufgabenforschung etablierten Analysekategorien. Durch die induktive Bildung von Subkategorien wurden die Hauptkategorien am Datenmaterial differenziert.²⁵ Gleichzeitig lässt es die deduktiv-induktive Mischform der Kategorienbildung zu, dass unerwartete Inhalte zur Bildung neuer Kategorien oder Subkategorien führen. Die Kategorien bestehen – angelehnt an KUCKARTZ Empfehlungen zur Kategorienbildung – aus einer Bezeichnung und einer inhaltlichen Beschreibung der Kategorie.²⁶ Darüber hinaus werden Subkategorien bzw. Merkmalsausprägungen benannt, welche durch Kodierregeln von anderen Merkmalsausprägungen abgegrenzt und durch ein Ankerbeispiel illustriert werden.²⁷ Die entstehenden Kategorien und Subkategorien erheben den Anspruch, disjunkt (trennscharf) und erschöpfend zu sein.²⁸ Sie strukturieren das Material, so dass auch das entstandene Kategoriensystem als ein Ergebnis der Untersuchung gelten kann.

5.2.1 Definition der Auswahl- und Analyseeinheit sowie des Aufgabenbegriffs

Alle im Untersuchungszeitraum gestellten Aufgabensets der drei Bundesländer Bayern, Niedersachsen und Thüringen dienen als Auswahlseinheiten der qualitativen Inhaltsanalyse. Auswahlseinheiten stellen die Auswahl aus der Gesamtheit aller potentiellen Untersuchungsobjekte dar, die der Analyse zugrunde liegen.²⁹ Als Aufgabensets werden in der Untersuchung die Gesamtheit der Aufgabenstellungen benannt, welche eine Schülerin oder ein Schüler im Rahmen einer schriftlichen Abiturprüfung zu lösen hat. Den Schülerinnen und Schülern stehen in Niedersachsen und Thüringen zwei, in Bayern vier Aufgabensets zur Verfügung, aus denen sie sich eines zur Bearbeitung auswählen.³⁰

²⁵ Vgl. a. a. O., 95.

²⁶ A. a. O., 67.

²⁷ Der Begriff des Ankerbeispiels stammt vom MAYRING, welcher ein ähnliches Vorgehen zur Kategorienbildung vorschlägt (vgl. MAYRING, Qualitative Inhaltsanalyse, 97).

²⁸ Vgl. KUCKARTZ, Qualitative Inhaltsanalyse, 67.

²⁹ Zum Begriff der Auswahlseinheiten vgl. a. a. O., 30.

³⁰ Die EPA sprechen in diesem Zusammenhang von einer Prüfungsaufgabe, welche sich aus mehreren Teilaufgaben zusammensetzt (vgl. EPA, 18). Da aus dieser Bezeichnung jedoch nicht hervor geht, dass es sich um mehrere Handlungsaufforderungen handelt, welche einen zusammengehörigen Aufgabenkomplex ergeben, wird der Begriff »Aufgabenset« dem der »Prüfungsaufgabe« vorgezogen.

Als Analyseeinheit der Studie dienen zum einen die Aufgabenstellungen und zum anderen die dazugehörigen Erwartungshorizonte.³¹ Der gesamten Analyse liegt anknüpfend an den aktuellen Forschungsstand der Aufgabenforschung ein enger Aufgabenbegriff zugrunde, welcher die Handlungsaufforderung als konstitutiv für eine Aufgabe bestimmt und somit als Aufgabe die kleinste zusammenhängende inhaltliche Anforderung definiert.³² Daraus folgt, dass die Analyseeinheiten eng an die Operatoren geknüpft sind und unabhängig von der Nummerierung der Abituraufgabensets erfasst werden. Wenn eine Aufgabenformulierung mehrere Operatoren und damit mehrere Handlungsaufforderungen besitzt, werden diese getrennt voneinander erfasst, auch wenn die Aufgaben inhaltlich aufeinander aufbauen. Auf diese Weise wird den Unterschieden zwischen den Bundesländern in der Aufgabengestaltung Rechnung getragen und die vom Prüfling zu erfüllenden Handlungsschritte werden vergleichbar. Ein Beispiel für zwei Handlungsanweisungen, welche als zwei Aufgaben getrennt voneinander erfasst werden, stellt folgende Aufgabenformulierung dar:

»Entfalten Sie verschiedene Beiträge zur Theodizeefrage und setzen Sie diese in Beziehung zur Aussage von Samuel Koch!« [BY, 2014 (2)]

Das Entfalten verschiedener Beiträge zur Theodizeefrage ist zwar eine Voraussetzung, um diese mit der Aussage Samuel Kochs in Beziehung zu setzen, es handelt sich jedoch um zwei Arbeitsgänge, die die Prüflinge zu bewerkstelligen haben.

Einen Ausnahmefall dieser Regel bilden die Handlungsanweisung »beziehen Sie dabei... ein« »ziehen Sie dabei ... heran«, »berücksichtigen Sie dabei...«. Es handelt sich zwar um Handlungsanweisungen, diese sind aber unmittelbar mit einer anderen Handlungsanweisung verknüpft und können nur gemeinsam in einem Arbeitsgang erfüllt werden. Dies ist beispielsweise in folgender Aufgabe der Fall:

»4. Nehmen Sie in einem Brief an den Vorstand der ACK Bayern einen begründeten Standpunkt zu deren Aufruf ein (M2). Beziehen Sie in Ihre Ausführungen ausgewählte Aspekte aus dem Text von Petra Bahr ein.« [NI, 2015, eA (2)]³³

³¹ Vgl. zum Begriff der Analyseeinheit KUCKARTZ, Qualitative Inhaltsanalyse, 30f.

³² Vgl. ALEXANDER JORDAN/NATHALIE ROSS/STEFAN KRAUSS/JÜRGEN BAUMERT/WERNER BLUM/MICHAEL NEUBRANDT/KATRIN LÖWEN/MARTIN BRUNNER/MAREIKE KUNTER, Klassifikationsschema für Mathematikaufgaben. Dokumentation der Aufgabenkategorisierung im COACTIV-Projekt. Berlin 2006. Sowie KÜHN: Steuerung und Innovation durch Abschlussprüfungen?, 139 und OTTO, »Interpretieren Sie das Gedicht...«, 145.

³³ Der Aufgabenstellung liegt ein Aufruf der Arbeitsgemeinschaft christlicher Kirchen Bayern zugrunde, sich in konkreten Handlungen für Geflüchtete zu engagieren.

Die Aufgabenstellung wird als *eine* Aufgabe gezählt.

Andersherum gibt es jedoch auch Aufgabenstellungen, die nur eine Handlungsanweisung formulieren, hinter welcher sich jedoch mehrere Aufgaben verbergen.³⁴ Ein Beispiel ist folgende Aufgabenstellung:

»Setzen Sie die Argumentation Kants in Beziehung zum christlichen Schöpfungsglaubens sowie der Reich-Gottes-Botschaft.« [TH, 2018, eA (2)]

Sowohl der christliche Schöpfungsglaube als auch die Reich-Gottes-Botschaft soll mit einem Auszug aus IMMANUEL KANTS Schrift *Kritik der reinen Vernunft* in Beziehung gesetzt werden. Da es sich um zwei verschiedene Modelle handelt, werden sie als zwei Aufgaben gezählt, obwohl nur ein Operator genannt wird.

Während alle im Untersuchungszeitraum gestellten Aufgaben Analyseeinheiten darstellen, werden als Fälle der vorliegenden Studie nur jene Aufgabenstellungen bezeichnet, welche Urteilsfähigkeit prüfen.³⁵ Die Anzahl der im Untersuchungszeitraum vorkommenden Fälle wird im Bereich quantitativer Aspekte erhoben (vgl. Kapitel 6.1). Die Gesamtzahl der in der Studie untersuchten Fälle wird als Untersuchungskorpus bezeichnet.³⁶

Neben den Aufgabenstellungen dienen ebenso die im Untersuchungszeitraum erstellten Erwartungshorizonte als Analyseeinheiten. Der Begriff *Erwartungshorizont* wird in der Studie einheitlich für die Bewertungshinweise für Lehrerinnen und Lehrer gebraucht, was der niedersächsischen Bezeichnung entspricht, während die Bewertungshinweise in Bayern offiziell als *Hinweise zur Korrektur und Bewertung der Abiturprüfungsaufgaben* und in Thüringen offiziell als *Hinweise für die Lehrerinnen und Lehrer* bezeichnet werden.

Die Aufgabenstellungen und Erwartungshorizonte werden mit einer Signatur versehen, um sie jederzeit zum jeweiligen Bundesland und zu ihrem ursprünglichen Aufgabenset zuordnen zu können. Die Signaturen der Aufgabenstellungen folgen dem Schema: [Kürzel des Bundeslands, Jahr, ggf. Anforderungsniveau³⁷ (Wahlaufgabenset)]. So ergibt sich für eine Aufgabe, welche

³⁴ Dieser Fall taucht in dem für diese Studie zugrunde liegenden Untersuchungsmaterial zweimal auf.

³⁵ Zum Begriff des Falls in qualitativen Studien vgl. KUCKARTZ, Qualitative Inhaltsanalyse, 49.

³⁶ Auf die Verwendung des in der qualitativen Forschung etablierten Begriffs Sample (bzw. Stichprobe) wurde bewusst verzichtet, da er in der Regel mit Probanden in Verbindung steht und daher in Bezug auf das Forschungsdesign dieser Studie zu Unklarheiten führen kann.

³⁷ Unter den untersuchten Bundesländern wurden allein in Niedersachsen schriftliche Abituraufgaben auf erhöhtem und grundlegendem Anforderungsniveau gestellt. In Thüringen wird ausschließlich auf erhöhtem Anforderungsniveau geprüft. In Bayern findet

2015 in Niedersachsen mit grundlegendem Anforderungsniveau als eine Aufgabe aus dem Wahlaufgabenset 2 gestellt wurde, folgende Signatur: [NI, 2015, gA (2)]. Der zur jeweiligen Aufgabenstellung gehörende Erwartungshorizont ist mit einem E und einem Bindestrich vor dem Länderkürzel gekennzeichnet, so dass sich für den Erwartungshorizont des vorliegenden Beispiels folgende Signatur ergibt: [E-NI, 2015, gA (2)].

5.2.2 Auswahl der Aufgabenstellungen zum Kompetenzbereich ethische und religiöse Urteilsfähigkeit

Der Kompetenzbereich Urteilsfähigkeit, der in dieser Untersuchung in den Fokus rückt, wird in den EPA mit der Anforderung konkretisiert, dass Schülerinnen und Schüler »in ethischen und religiösen Fragen begründet urteilen« können.³⁸ Bezugnehmend auf diese Anforderung und vor dem Hintergrund der fachlichen Grundlegungen in Kapitel 2.1 wird bestimmt, dass als Fall jene Aufgaben in die Studie aufgenommen werden, die folgender Definition entsprechen:

Eine Aufgabenstellung wird dann in das Untersuchungsmaterial aufgenommen, wenn in ihr eine eigenständige Position im Sinne von Zustimmung oder Ablehnung hinsichtlich einer kontroversen ethischen oder religiösen Fragestellung vertreten werden soll.

»Kontrovers« bedeutet in der Definition, dass mindestens zwei konträre, vertretbare Positionen in Bezug auf die Fragestellung gefunden werden können. Dies ist konstitutiv für Urteilsbildungsprozesse, da ein Urteilsentscheid nur notwendig wird, wenn es mehrere plausibel zu begründende und sich widersprechende Optionen gibt, wobei letzteres die Möglichkeit einer Kompromisslösung nicht ausschließt. Es handelt sich um Aufgabenstellung, die eine Wertung beinhalten (Werturteile) und daher mit den Einstellungen und Haltungen der Prüflinge verknüpft sind (vgl. Kapitel 2.1).

Um die Eigenständigkeit der Position sicher zu stellen, muss die Aufgabe hinsichtlich ihrer Lösung offen sein (vgl. ebenso Kapitel 2.1), das heißt, die möglichen konträren Positionen müssen gleichermaßen zu einer sehr guten Bewertung der Lösung führen können. Der Prüfling muss sich sowohl für eine Zustimmung als auch für eine Ablehnung entscheiden können.

In den EPA werden die für die Aufgabenstellungen zu nutzenden Operatoren inhaltlich definiert (vgl. Kap. 3.2.2). Ausgehend von den in den EPA gesetzten Vorgaben kommen Aufgaben mit folgenden Operatoren für den Kompetenzbe-

keine Differenzierung zwischen erhöhtem und grundlegendem Anforderungsniveau bzw. in Leistungs- und Grundkurs statt.

³⁸ Vgl. EPA, 8.

reich Urteilsfähigkeit in Frage: »sich auseinandersetzen mit...«, »beurteilen«, »bewerten«, »Stellung nehmen«, »einen begründeten Standpunkt einnehmen«, »erörtern«, »prüfen«, »überprüfen«, »einen begründeten Standpunkt einnehmen«.³⁹

5.2.3 Grenz- und Sonderfälle

Das zunächst erprobte Vorgehen, die Aufgaben zur Urteilsfähigkeit anhand der Operatoren der Aufgabe herauszufiltern, hat sich als nicht gangbar erwiesen, da im Datenmaterial auch Aufgabenstellungen vorliegen, welche eine eigenständige Positionierung verlangen, ohne einen entsprechenden Operator zu nutzen (vgl. Kapitel 5.4.2). Das Kriterium der Zuordnung stellt daher explizit nicht der Operator einer Aufgabe dar, sondern in erster Linie die oben genannte inhaltliche Definition. So werden beispielsweise Gestaltungsaufgaben mit den Operatoren »gestalten« oder »entwerfen«, in denen das Endprodukt eine Urteilsbildung enthalten soll, in die Studie aufgenommen, obwohl sie berechtigter Weise keinen der oben genannten Operatoren benutzen. Gleichzeitig gibt es Aufgabenstellungen mit Operatoren der Beurteilung, die jedoch keine kontroverse Fragestellung aufwerfen, sondern ein reines Sachurteil fordern und daher aus dem Untersuchungsmaterial ausgeschlossen werden. Neben diesen inhaltlich gut nachvollziehbaren Fälle treten Aufgaben, in denen der Anforderungsbereich (AFB) des Operators nicht der inhaltlichen Anforderung der Aufgabe entspricht. Beispiele für die verschiedenen Grenz- und Sonderfälle werden im Folgenden kurz beleuchtet.

a) Urteilsbildungen mit Perspektivenwechsel

Unter den 290 Aufgaben der Auswahlinheit befinden sich neun Aufgabenstellungen, die eine Urteilsbildung aus einer vorgegebenen Perspektive fordern (acht aus Bayern, eine aus Thüringen). Die einzunehmende Perspektive ist in allen neun Fällen eine christliche bzw. evangelische (vgl. Anhang IV). Diese Aufgabenstellungen sind durch die oben genannte Definition vom Untersuchungsmaterial ausgeschlossen, da die Eigenständigkeit der Positionierung nicht gegeben ist. Die persönlichen Einstellungen und Haltungen der Schülerinnen und Schüler spielen zur Beantwortung der Fragen keine Rolle. Die Ergebnisse können in der Regel eindeutig in den Kategorien »richtig« und »falsch« gefasst werden.

In der Operatorenliste der EPA werden die zum Perspektivwechsel gehörenden Operatoren, »Stellung nehmen aus der Sicht von...« sowie »eine Erwiderung formulieren aus der Sicht von...« dem AFB III zugeordnet.⁴⁰ Ein Beispiel –

³⁹ Vgl. Operatorenliste a.a.O., 14. Siehe auch Anhang I.

⁴⁰ Vgl. EPA, 14.

mit leicht abweichender Operatorenformulierung – gibt folgende Aufgabenstellung:

»4. In einer Online-Debatte darüber, was nach dem Tod mit Profilen und Konten bei sozialen Netzwerken im Internet geschehen soll, findet sich folgender Beitrag:

«So lange dein Profil aktiv ist, wirst du niemals sterben.»

[...]

4.2 Beurteilen Sie diese Vorstellung aus christlicher Perspektive! Gehen Sie dabei auch auf neutestamentliche Aussagen ein!« [BY, 2015 (1)]

Da ein Plural an christlichen Perspektiven existiert, sind verschiedene Argumentationen möglich. Dennoch lässt sich fragen, ob die Anforderung einer Perspektivübernahme nicht eher eine Transferleistung darstellt, welche dem AFB II zugeordnet werden müsste, anstatt zur Kategorie der eigenständigen Problemlösungen (AFB III): So müssen die Prüflinge im vorliegenden Beispiel die christliche Lehre *kennen* und sie auf die Frage nach einem Weiterleben nach dem Tod durch mediale Möglichkeiten *übertragen*. Zwei Wissensbereiche sind zu verknüpfen, ohne dass eigenständige Deutungen zwingend notwendig werden.

b) Der Operator »erörtern«

Für die der Studie zugrunde liegende Definition von Urteilsfähigkeit ist die eigenständige Positionierung in einer Aufgabenstellung konstitutiv. In Bezug auf den Operator »erörtern«, welcher im Zusammenhang mit einer kontroversen Fragestellung zweimal in der Auswahlinheit vorkommt, ist eine Positionierung nicht explizit angesprochen, wie an folgendem Beispiel illustriert werden soll:

»Die Achtung vor der Wahrheit, die Wahrung der Menschenwürde und die wahrhaftige Unterrichtung der Öffentlichkeit sind oberste Gebote der Presse.« (Ziffer 1, Pressekodex des Deutschen Presserates)

Erörtern Sie Möglichkeiten und Grenzen der Umsetzung dieses Grundsatzes unter Berücksichtigung der im Text (Seite 2 f.) entfalteten Gedanken! Beziehen Sie dabei grundlegende Aspekte philosophischer und christlicher Ethik mit ein!« [BY 2017 (1)]

Die Aufforderung »Möglichkeiten und Grenzen« aufzuzeigen, lässt auf die Beachtung von Pro- und Kontraargumenten in Hinblick auf die Möglichkeiten der Umsetzung des Pressekodex schließen. Die Aufgabe fordert folglich die im Forschungsdiskurs wiederholt als Qualitätsmerkmale genannte Mehrperspektivität ein. Dass eine persönliche Positionierung zu einer der Optionen stattfinden soll, ist nicht explizit erwähnt, wodurch die Zuordnung uneindeutig ist. In den EPA wird der Operator »erörtern« folgendermaßen definiert: »die Vielschichtigkeit eines Beurteilungsproblems erkennen und darstellen, dazu Thesen erfassen

bzw. aufstellen, Argumente formulieren, nachvollziehbare Zusammenhänge herstellen und dabei eine begründete Schlussfolgerung erarbeiten (dialektische Erörterung)«. ⁴¹ Der Hinweis auf die begründete Schlussfolgerung am Ende der Überlegungen lässt eine Positionierung vermuten. Ebenso kann bei der oben zitierten Aufgabenstellung angenommen werden, dass dieser die Aufforderung zur Urteilsbildung implizit zugrunde liegt. Die beiden Aufgabenstellungen mit dem Operator »erörtern« werden daher als Fälle in die Studie aufgenommen ([BY 2017 (1)] und [TH, 2017, eA (2)]). ⁴²

c) Die Operatoren »Konsequenzen aufzeigen« und »Perspektiven entwickeln«

Eine weitere Gruppe der Grenzfälle stellen Aufgabenstellungen mit den Operatoren »Konsequenzen aufzeigen« und »Perspektiven entwickeln« dar. Die Auswahlinheit enthält zwölf Aufgabenstellungen dieser Art (neun aus Bayern, drei aus Niedersachsen), wobei diese mit nur einer Ausnahme stets mit einer christlichen Sicht verbunden sind, also christliche Perspektiven aufgezeigt oder die Konsequenzen eines christlichen Verständnisses dargestellt werden sollen. Beispiel geben folgende Aufgabenstellung:

»5.2 Entwickeln Sie an einer konkreten ethischen Fragestellung Perspektiven christlichen Handelns, die die Gedanken Leonardo Boffs aufnehmen!« ⁴³ [BY, 2015 (2)]
 »4 »Die Bilder, die Menschen von Gott haben, werden nicht ohne Auswirkungen darauf bleiben, wie sie ihr Verhalten und ihre Beziehungen zu den Mitmenschen gestalten.« (Zeile 28–31) Entfalten Sie eine frei gewählte ethische Fragestellung und zeigen Sie Konsequenzen christlichen Gottesverständnisses für den Umgang mit dieser Fragestellung auf!« ⁴⁴ [BY, 2017 (2)]

Diese Aufgaben erfordern laut EPA, »Schlussfolgerungen [zu] ziehen [und] Perspektiven, Modelle, Handlungsmöglichkeiten, Konzepte u. a. [zu] entfalten« ⁴⁵. Die Forderung einer eigenständigen Urteilsbildung wird nicht erwähnt. Erwartet

⁴¹ Vgl. ebd.

⁴² Für die Aufnahme des Operators »erörtern« spricht zudem, dass die »Erörterung« in der Beschreibung des AFB III mit einer eigenständigen Urteilsbildung in Zusammenhang gebracht wird. Dort heißt es: »Erörtern von fachspezifischen Positionen, Thesen und Problemen mit dem Ziel einer begründeten und überzeugenden Stellungnahme« (a. a. O., 12).

⁴³ Der Aufgabenstellung liegt ein Auszug aus LEONARD BOFFS »Kleine[n] Trinitätslehre« zugrunde, welcher den Schülerinnen und Schülern vorliegt.

⁴⁴ Der Aufgabenstellung liegt eine Rede des früheren EKD-Ratsvorsitzenden NIKOLAUS SCHNEIDER zugrunde, aus welcher auch das in der Aufgabe enthaltene Zitat entnommen ist und welche den Prüflingen vorlag.

⁴⁵ Vgl. EPA, 15.

wird, dass der Prüfling einen Zielzustand definiert und Handlungsmöglichkeiten aufzeigt, die diesen Zielzustand fördern würden. Ähnlichkeit haben diese Aufgaben mit den sogenannten Gestaltungsurteilen, welche von JOACHIM DETJEN et al. in Bezug auf Politische Urteilsfähigkeit als eine Urteilsart neben dem Feststellungs-, Erweiterungs-, Wert- und Entscheidungsurteil kategorisiert werden.⁴⁶ Gestaltungsurteile sind dort als inhaltliche Lösungen sachlicher Probleme definiert, in denen »konkrete Angaben über eine zu regelnde Materie« gemacht werden sollen.⁴⁷ Trotz dieser Ähnlichkeit bleiben die Aufgabenstellungen mit den Operatoren »Konsequenzen aufzeigen« und »Perspektiven entwickeln« hinsichtlich der Haltung der Prüflinge offen. Da Modelle, Handlungsmöglichkeiten und Konzepte oder auch eine christliche Perspektive auf einen spezifischen Fall angewendet werden sollen, handelt es sich bei diesen Aufgabenstellungen eher um Transferleistungen (AFB II), so dass sie nicht dem hier untersuchten Kompetenzbereich Urteilsfähigkeit zugeordnet werden.

d) Operatoren aus dem Anforderungsbereich I und II

Unter den 290 im Untersuchungszeitraum formulierten Aufgabenstellungen lassen sich darüber hinaus auch solche finden, welche Operatoren aus dem AFB II nutzen (»Belegen«, »Nachweisen«, »Begründen«, »Erläutern«, »Erklären«, »Entfalten«), inhaltlich aber eine Positionierung fordern, die nicht für den AFB II vorgesehen ist.⁴⁸ Mit Blick auf die genannten Operatoren ist für die Zuordnung entscheidend, ob es um kontroverse ethische Thesen geht, welche belegt, begründet oder entfaltet werden sollen – dann handelt es sich um eine Fähigkeit, die dem Kompetenzbereich Urteilsfähigkeit zuzurechnen ist, oder ob es sich um einen reinen Sachverhalt handelt, der nachgewiesen werden soll – in diesem Fall ist die Aufgabe nicht dem Kompetenzbereich Urteilsfähigkeit zuzuordnen, sondern eine reine Transferleistung aus AFB II. Ein Beispiel für eine Aufgabenstellung, die eine kontroverse Fragestellung aufgreift, eine Positionierung fordert und daher dem Kompetenzbereich Urteilsfähigkeit zuzuordnen ist, lautet folgendermaßen:

⁴⁶ Vgl. JOACHIM DETJEN/PETER MASSING/DAGMAR RICHTER/GEORG WEIBENO, Politikkompetenz – ein Modell. Wiesbaden 2012, 53. Die von DETJEN et al. verwendete Bezeichnung »Gestaltungsurteil« lässt eine Nähe zu dem in den EPA benannten Kompetenzbereich der *Gestaltungsfähigkeit* vermuten. Dieser zielt jedoch vor allem auf eine mediale Ausdrucks- und Gestaltungsform im Sinne von Textgestaltungen oder -transformationen. Gesellschaftliche Gestaltungsprozesse sind in dem Kompetenzbereich Gestaltungsfähigkeit nicht bedacht, so dass die Zuordnung der Aufgaben in den Kompetenzbereich *Urteilsfähigkeit* plausibler erscheint (vgl. EPA, 9).

⁴⁷ Vgl. JOACHIM DETJEN et al., Politikkompetenz, 35 sowie 57.

⁴⁸ Vgl. EPA, 13.

»5. Dankbarkeit und Demut (vgl. Material 1, Z. 18–20) auch im Leid? Entfalten Sie dazu eine Antwort und greifen Sie dabei auf einen Ansatz zur Theodizee zurück!« [BY, 2015 (3)]

Der Prüfling wird aufgefordert, zu einer Haltung der Dankbarkeit in Leidsituationen Stellung zu nehmen. Er kann diese Haltung ablehnen oder ihr zustimmen.

Einen hervorzuhebenden Sonderfall, der *in das Datenmaterial aufgenommen* wurde, stellt in diesem Zusammenhang die zweite Aufgabe des folgenden Aufgabensets dar:

»Entwerfen Sie auf der Grundlage der Materialien einen Essay, in dem Sie sich mit der Botschaft Jesu auseinandersetzen.

Zeigen Sie auf, welche Rolle christliche Maßstäbe für Ihre Entscheidungen spielen.« [TH, 2014 (2)].

Der Operator »Aufzeigen« ist dem AFB I zuzuordnen und verlangt laut seiner Definition in den EPA »den Gedankengang oder die Hauptaussage eines Textes oder einer Position mit eigenen Worten dar[zu]legen«.⁴⁹ Entsprechend der Logik der AFB wird mit diesem Operator eine Reproduktion und keine eigenständige Urteilsbildung erwartet. Inhaltlich erfordert diese Aufgabe jedoch, dass der Prüfling seine eigene Moral reflektiert und christliche Maßstäbe hinsichtlich ihrer Relevanz für das eigene Handeln beurteilt. Er muss sich also durchaus positionieren. Der Operator wurde an dieser Stelle falsch gewählt.

Anhand der vorgestellten Fälle wird nochmals explizit deutlich, dass sich eine an den Operatoren ausgerichtete Auswahl der Fragestellungen, die eine Urteilsbildung fordern, als weniger sinnvoll erweist als eine inhaltlich ausgerichtete Definition von Urteilsbildungsaufgaben, anhand derer die Auswahl erfolgt.

Die Kategorie *Passung zwischen geforderter und erwarteter Urteilsfähigkeit* trägt dem Aufkommen dieser Unstimmigkeiten Rechnung und untersucht nochmals die Stimmigkeit der genutzten Operatoren in Bezug zu den Anforderungen der Erwartungshorizonte (vgl. Kapitel 6.3.1).

Nicht in das zu analysierende Datenmaterial aufgenommen wurden trotz Bezug zur Ethik aufgrund der fehlenden Positionierungsforderung hingegen beispielsweise folgende Aufgabenstellungen:

»2. Setzen Sie die im Text beschriebenen Phänomene der Selbstinszenierung in einer Casting-Gesellschaft und evangelisches Menschenverständnis in Beziehung zueinander!« [BY, 2015 (1)]

⁴⁹ Vgl. ebd.

»5 Entfalten Sie Gründe und Konkretionen eines empathischen Umgangs mit Flüchtlingen aus der Sicht eines Ansatzes philosophischer Ethik!« [BY, 2016 (3)]

In beiden Aufgabenstellungen ist die persönliche Haltung des Prüflings insofern irrelevant, als konkrete Konzeptionen auf eine Fragestellung angewendet werden sollen. Was der Prüfling über die Selbstinszenierung in der Casting-Gesellschaft oder über einen empathischen Umgang mit Geflüchteten denkt, wird nicht explizit erfragt. Ob implizit diese Anforderungen mitschwingen, wird bei der Analyse der Erwartungshorizonte in der Kategorie *Passung zwischen geforderter und erwarteter Urteilsfähigkeit* erhoben, welche alle Erwartungshorizonte in den Blick nimmt (vgl. Kapitel 6.3.1).

e) Die Operatoren »entwerfen« und »gestalten«

Ein weiterer zu besprechender »Sonderfall« besteht in Bezug auf die Operatoren »entwerfen« und »gestalten«. Sie werden in den EPA im AFB III folgendermaßen konkretisiert: »sich textbezogen mit einer Fragestellung kreativ auseinander setzen [sic]«. ⁵⁰ Je nachdem, was den Gegenstand der kreativen Auseinandersetzung darstellt bzw. was entworfen oder gestaltet werden soll (Leserbrief, Forenbeitrag für das Internet, eine Rede etc.) können Aufgabenstellungen mit diesem Operator eine Positionierung in Bezug auf eine ethische Fragestellung fordern oder auch nicht. Daran wird deutlich, dass die Aufgaben dieses Typs nicht ganz eindeutig hinsichtlich der Notwendigkeit einer Positionierung sind. Wenn die Aufgabenstellung jedoch in dem zu entwickelnden Produkt eine Auseinandersetzung mit einer kontroversen Fragestellung fordert, kann die Aufgabe eindeutig in den Bereich der Urteilsfähigkeit eingeordnet werden, da der Operator »sich auseinandersetzen mit« ein »begründetes eigenes Urteil zu einer Position oder einem dargestellten Sachverhalt [...]« verlangt. ⁵¹ Aufgenommen wurde beispielsweise die dritte Handlungsanweisung der folgenden Aufgabenstellung:

»Entwerfen Sie einen offenen Brief, in dem Sie

- das in dem »Glaubensbekenntnis« des Textbeispiels enthaltene Gottesbild darstellen.
- dieses mit dem Ihnen bekannten biblischen Gottesbild vergleichen
- sich persönlich mit den Gottesbildern **auseinandersetzen**.« [TH, 2019, eA (1)]

Uneindeutig sind hingegen folgende zwei Beispiele, da Streitbar ist, ob ein Artikel in einer (Schüler-)Zeitung grundsätzlich einer Positionierung bedarf:

⁵⁰ Vgl. a. a. O., 14.

⁵¹ Vgl. ebd.

»Wir müssen lernen, damit umzugehen. Mit unserem eigenen Neid – und dem der anderen.« (Material 1, Zeile 57 f.)

Entwerfen Sie einen Artikel für die Schülerzeitung über den Umgang mit Neid im Schulleben! Beziehen Sie Gedanken der Materialien 1–4 und Aspekte biblischer Anthropologie und Ethik mit ein!« [BY, 2017 (3)].

»Die Redaktion der Schülerzeitung Ihrer Schule plant eine Themenausgabe mit dem Titel ›Leben in der Welt von morgen‹. Dabei sollen die unterschiedlichen Fächer zu Wort kommen.

Entwerfen Sie für das Fach Evangelische Religionslehre einen Beitrag, der sich mit Chancen und Grenzen des technisch Machbaren auseinandersetzt! Gehen Sie dabei von zentralen Gedanken des Textes auf Seite 12 f. aus, wählen Sie aber eine andere Thematik als die des autonomen Fahrens! **Formulieren Sie** für Ihren Textbeitrag eine passende Überschrift!« [BY, 2018 (4)]

Die Aufgaben werden in das Datenmaterial aufgenommen, da es sich um ethische Fragestellungen handelt – den Umgang mit Neid und die Chancen und Grenzen des technisch Machbaren – welche normativer Prämissen zur Beantwortung bedürfen (vgl. Kapitel 2.1). In Bezug auf letztere Aufgabe kann diskutiert werden, ob es sich um eine Aufforderung zur Perspektivübernahme handelt, da der Schülerzeitungsartikel für das Fach Evangelische Religionslehre geschrieben werden soll. Allerdings gibt diese Vorgabe ausschließlich den *Kontext* vor, in dem argumentiert werden soll, jedoch nicht die Richtung des Urteilsentscheidendes. Es kann sowohl für eine intensive Nutzung allen technisch Machbaren argumentiert werden als auch gegen diese. Die Urteilsbildung ist damit *kontrovers* hinsichtlich ihrer Fragestellung und *offen* hinsichtlich der Urteilsbildung.

f) Fragestellung, die weder als »ethisch« noch als »religiös« zu klassifizieren sind

Aus dem Datenmaterial ausgeschlossen wurden Aufgabenstellungen, welche ein Urteil über ein Fragestellungen fordern, die weder ethischer noch religiöser Art sind. Ein Beispiel stellt folgende Aufgabenstellung aus Bayern dar, in der zwei Bilder zur Trinität hinsichtlich ihrer Passung zum trinitarischen Gottesverständnisses des brasilianischen, katholischen Theologen LEONARDO BOFFS bewertet werden sollen.

»3 **Begründen Sie**, welches Kunstwerk (Material 2 oder Material 3) Ihrer Ansicht nach besser geeignet ist, die zentralen Aussagen des Textes zu veranschaulichen!« [BY, 2015 (2)]

Da die Aufgabenstellung keine Stellungnahme des Prüflings zur Trinität oder zu den Gedanken BOFFS fordert, sondern ein Urteil über die Passung zwischen Text und Bild, ist die Aufgabenstellung nur in geringem Maße an die ethischen oder

religiösen Einstellungen und Haltungen der Prüflinge geknüpft und wird daher aufgrund objektiver Bewertbarkeit nicht als Fall in die Studie aufgenommen.⁵² Ähnlich verhält es sich mit folgender Aufgabenstellung aus Niedersachsen, in der die Passung zwischen einem Poster der Initiative »Brot für die Welt« und den theologischen Aussagen des evangelischen Theologen HELMUT FISCHER zum Abendmahl bewertet werden soll.

»4 Überprüfen Sie, ob das Poster der Initiative »Brot für die Welt« (M2) eine geeignete Veranschaulichung der Inhalte des Textes von Helmut Fischer darstellt.«⁵³ [NI 2014 eA (2)]

Auch in dieser Aufgabe steht nicht die Interpretation des Abendmahls durch das Poster oder durch den Text FISCHERS zu Diskussion, sondern lediglich die Passung beider, so dass auch diese Aufgabenstellung aus dem Analysematerial ausgeschlossen wird.

g) Operator »interpretieren«

Einen letzten zu besprechenden Sonderfall stellt eine Aufgabenstellung aus Thüringen dar, bei der der Operator »interpretieren« mit einer ethischen Fragestellung nach der Gestaltung von Freiheit im Leben verbunden wird. Der Operator »interpretieren« fordert laut den EPA »einen Text oder ein anderes Material (zum Beispiel Bild, Karikatur, Tondokument, Film) sachgemäß zu analysieren und auf der Basis methodisch reflektierten Deutens zu einer schlüssigen Gesamtauslegung [zu] gelangen«⁵⁴. Auch wenn es sich um einen Operator des AFB III handelt, ist eine Positionierung nicht vorgesehen. Anders verhält es sich im Kontext folgender Aufgabe:

»[3] Interpretieren Sie – auch mit Hilfe Ihrer Kenntnisse vom christlichen Menschenbild und der evangelischen Sozialethik – das Gedicht »Stufen« von Hermann Hesse (Material 2) unter der Fragestellung, wie sich Freiheit im Leben verantwortlich gestalten lässt.« [TH, 2017, eA (1)].

⁵² Darüber hinaus bleibt der Operator »begründen« aus dem AFB II, welcher definiert ist als »Aussagen mit Argumenten stützen«, uneindeutig hinsichtlich einer geforderten Positionierung, wobei auf inhaltlicher Ebene der Aufgabe deutlich wird, dass vom Prüfling ein Urteilsentscheid verlangt wird.

⁵³ Das angesprochene Poster bildet das Abendmahl Jesu in zweifacher Ausführung ab: In der ersten sitzt ein weißer Mann in der Mitte der Tafel vor übertollen Tellern, während die Teller der am Rand der Tafel Sitzenden leer sind. In der zweiten Variante sind die Speisen auf die am Tisch Sitzenden verteilt. Der weiße Mann hat seine zentrale Sitzposition verlassen.

⁵⁴ EPA, 16.

Durch die kontroverse Fragestellung nach dem verantwortungsvollen Umgang mit der Freiheit im Leben, wird implizit eine Urteilsbildung gefordert, so dass die Aufgabenstellung in das Datenmaterial aufgenommen wird.

Resümierend kann festgehalten werden, dass durch die benannten Grenz- und Sonderfälle offensichtlich wird, dass eine unmittelbare Erfassung, ob eine eigenständige Urteilsbildung gefordert ist, nicht immer möglich ist. Das entscheidende Argument für die Auswahl bildet nicht der verwendete Operator, sondern die Anforderungen, welche mit einer Aufgabenstellung verbunden sind. Die bereits erwähnte Kategorie *Passung zwischen geforderter und erwarteter Urteilsfähigkeit* nimmt auftretende Uneindeutigkeiten nochmals in den Blick (vgl. Kapitel 6.3.1).

5.3 Kategoriensystem der Studie

Nachdem im Vorangegangenen deutlich wurde, anhand welcher Kriterien die Fälle der Studie aus dem Datensatz entnommen und wie mit Sonderfällen umgegangen wurde, stellt dieses Kapitel das Kategoriensystem vor, welches der Analyse der vorher extrahierten Aufgabenstellungen zur Urteilsfähigkeit dient.

Für eine klare Systematisierung werden die Kategorien des erstellten Kategoriensystems drei Gruppen zugeordnet: Die erste Kategoriengruppe erfasst allgemeine, *quantitative Aspekte* der Abituraufgabenstellungen, die zweite Gruppe erfasst *Merkmale der Aufgabenstellungen*, während eine dritte Gruppe auf die *Merkmale der Erwartungshorizonte* fokussiert ist (vgl. Tab. 13, siehe unten).

Mittels der Kategorien über *quantitative Aspekte* wird beantwortet, wie hoch der Anteil eigenständiger Urteilsbildungen in kontroversen ethischen und religiösen Fragestellungen ist und ob es in Bezug auf die untersuchten Bundesländer Unterschiede in diesem Anteil gibt. Dazu wird die *Anzahl der Aufgaben je Aufgabenset* im Allgemeinen sowie die Anzahl der Aufgaben, welche auf den Kompetenzbereich Urteilsfähigkeit entfallen, erfasst. Der Anteil, der Urteilsfähigkeit in Bezug auf die Gesamtzahl der Aufgaben ausmacht, ist nicht mit dem Anteil des AFB III gleichzusetzen, welcher laut EPA 30 % betragen soll. Grund dafür ist, dass in AFB III neben Aufgaben zur Urteilsbildung auch beispielsweise Gestaltungs- und Interpretationsaufgaben gestellt werden, die keine eigenständige Positionierung hinsichtlich ethischer oder religiöser Fragestellungen fordern (vgl. Kapitel 3.2.2).

Die Kategorien, welche sich auf die *Aufgabenstellungen* beziehen, erfassen, auf welche Art und Weise die Anforderungen an Urteilsfähigkeit in ethischen und religiösen Fragen geprüft werden. Die Kategorie *Aufgabenart* legt die drei in den EPA benannten Aufgabenarten zugrunde und hat damit sowohl den Materialbezug als auch die Bearbeitungsform im Blick. Die Kategorie *Fragestellungen der Urteilsbildung* untersucht, anhand welcher inhaltlichen Fragestellungen im Fach

Evangelische Religionslehre geurteilt werden soll. Eine weitere Kategorie fragt, welche Reflexionsanforderungen die untersuchten Aufgabenstellungen an ein Urteilsentscheid stellen, während eine weitere Kategorie Auskunft über die Textsorten der den Aufgaben zugrunde liegenden Materialien gibt.

In Bezug auf die *Erwartungshorizonte* treten Bewertungsaspekte und Fragen der Vergleichbarkeit der Bewertungsvorgaben in den Fokus. In der Kategorie *Passung zwischen geforderter und erwarteter Urteilsfähigkeit* wird die Stimmigkeit der verwendeten Operatoren untersucht sowie das Vorhandensein von Bewertungshinweisen zur Urteilsfähigkeit in den Erwartungshorizonten. Sehr zentral ist eine Kategorie, welche die Qualitätsmerkmale für Urteilsbildung erfasst, indem die in den Erwartungshorizonten genannten Bewertungskriterien analysiert werden. Zwei weitere Kategorien analysieren, ob die Erwartungshorizonte explizite Hinweise zum Umgang mit Schülerpositionen enthalten und ob sie mindestens zwei mögliche, konträre Positionen zur Fragestellung formulieren.

Folgende Übersicht gibt einen Gesamtüberblick über das entwickelte Kategoriensystem:

quantitative Aspekte	
• Anzahl der Aufgaben je Aufgabenset • Anteil der Aufgaben, die eine eigenständige Urteilsbildung fordern	
Ebene der Aufgabenstellung	Ebene der Erwartungshorizonte
• A) Aufgabenart (in EPA vorgegeben) • Textaufgabe • erweiterte Textaufgabe • Gestaltungsaufgabe • B) Fragestellungen der Urteilsbildung • C) Reflexionskriterien • D) Textsorten zugrundeliegender Materialien	• E) Passung zwischen geforderter und erwarteter Urteilsfähigkeit • F) Qualitätsmerkmale für Urteilsbildung • G) explizite Hinweise zum Umgang mit Schülerpositionen • H) Formulierung konträrer Positionen

Tab. 13: Kategoriensystem zur Urteilsfähigkeit in kontroversen ethischen und religiösen Fragen im schriftlichen Abitur des Faches Evangelische Religionslehre.

Das Analysesystem bezieht sich ausschließlich auf die Aufgaben zum Kompetenzbereich Urteilsfähigkeit und ist inhaltlich an diesen angepasst. Eine Ausnahme bilden lediglich die Kategorien *Anzahl der Aufgaben je Aufgabenset* sowie *Passung zwischen geforderter und erwarteter Urteilsfähigkeit*, welche die Gesamtheit der im Untersuchungszeitraum gestellten Abituraufgaben im Blick haben.

Im Folgenden werden die Kategorien des Kategoriensystems nochmals genauer beschrieben, Merkmalsausprägungen benannt, Kodierregeln aufgestellt und mit Ankerbeispielen versehen.

5.3.1 Quantitative Aspekte

Anzahl der Aufgaben je Aufgabenset

Die Kategorie *Anzahl der Aufgabe je Aufgabenset* erhebt, wie viele Aufgaben – im Sinne des in Kapitel 5.2.1 definierten engen Aufgabenbegriffs, der eine Aufgabe als die kleinste abgrenzbare Handlungsaufforderung bestimmt – ein Aufgabenset enthält. Da die Anzahl der Aufgaben pro Aufgabenset in den drei untersuchten Bundesländern stark variieren, ist diese Kategorie für den Vergleich der Bundesländer von Bedeutung.

Anteil der Aufgabenstellungen, die eine eigenständige ethische Urteilsbildung fordern

In dieser Kategorie wird untersucht, wie hoch der prozentuale Anteil gemessen an der Gesamtaufgabenzahl der Aufgabestellungen ist, welche eine Urteilsbildung im Sinne der oben genannten Definition von Urteilsfähigkeit fordern.

5.3.2 Kategorien zur Analyse der Aufgabenstellungen

A) Aufgabenart

Die Merkmale der Kategorie *Aufgabenart* sind durch den Materialbezug einer Aufgabe und die Bearbeitungsform dieser bestimmt. In Anlehnung an die EPA wird innerhalb der Kategorie zwischen Textaufgabe, erweiterter Textaufgabe und Gestaltungsaufgabe unterschieden.⁵⁵ Die Bearbeitungsform einer Aufgabe kann von der Auseinandersetzung mit einem Text (Textaufgabe) über ein Ins-Verhältnis-Setzen von Text und anderem Material wie beispielsweise einem Bild (erweiterte Textaufgabe) bis hin zur produktorientierten Transformation von Materialien gehen (Gestaltungsaufgabe). Die Offenheit der Ergebnisse nimmt in der beschriebenen Reihenfolge der Aufgabenarten zu. Neben diesen drei den

⁵⁵ Vgl. EPA, 16–18.

EPA entnommenen Aufgabenarten erfasst eine vierte Merkmalsausprägung alle Aufgaben, welche keinen Materialbezug besitzen und daher weder eine Textaufgabe noch eine Gestaltungsaufgabe darstellen.⁵⁶ Gestaltungsaufgaben gelten als innovative Aufgabenformen, deren Implementierung in den Unterricht durch das Zentralabitur vorangetrieben werden soll.⁵⁷ Sie haben zum Merkmal, dass sie im Gegensatz zu Textaufgaben und erweiterten Textaufgaben in situative Kontexte eingebettet sind.

Kategorie A: Aufgabenart		
Kodierung	Merkmalsausprägung	Beschreibung
A0	Aufgaben ohne Materialbezug	Die Aufgabe muss ausschließlich mittels eigenen Vorwissens der Prüflinge gelöst werden, ohne dass dabei auf vorliegende Materialien zurückgegriffen werden kann.
A1	Textaufgabe	Der Aufgabe liegt ein Text zugrunde, welcher zur Lösung der Aufgabe erschlossen und bearbeitet werden soll, ohne dass dabei eine kontextualisierte Produktorientierung angestrebt wird, wie es bei Gestaltungsaufgaben der Fall ist.
A2	Erweiterte Textaufgabe	Die Aufgabe fordert die Auseinandersetzung mit einem Text und weiteren Materialien (Karikaturen, Grafiken, Plastiken, Malerei...).
A3	Gestaltungsaufgabe	Die Aufgabe verlangt eine adressatengebundene Produktorientierung (wie bspw. Leserbrief, Forenbeitrag im Internet etc.) unter Einbezug und/oder Bearbeitung von Materialien. ⁵⁸

⁵⁶ In den EPA wird zwar der Anspruch formuliert, dass sich »Abituraufgaben im Fach Evangelische Religionslehre [...] auf Materialien [stützen]« (EPA, 16). Diese Aussage kann jedoch auf das gesamte Aufgabenset bezogen sein, so dass einzelne Handlungsaufforderungen nicht zwingend Materialbezug aufweisen müssen.

⁵⁷ Vgl. OTTO/KÜHN, Zwischen Tradition und Innovation, 6.

⁵⁸ Entsprechend der EPA wird »[u]nter produktionsorientierter Lösung [...] die Herstellung eines adressatenbezogenen Textes verstanden, der die Beherrschung der formalen und inhaltlichen Kriterien der entsprechenden Textgattung voraussetzt.« (EPA, 17). Als Beispiele werden in den EPA genannt: »Leserbrief, Rezension, Interview, Zeitungsartikel, Kommentar, Glosse, Essay, Dialog, Rede, Brief, Gleichnis, Liedtext« (EPA, 18).

Ankerbeispiele Kategorie *Aufgabenart*:

Aufgabe ohne Materialbezug:

»5 Setzen Sie sich damit auseinander, inwieweit grundlegende Aussagen des Apostolischen Glaubensbekenntnisses »die letzten Fragen nach dem Sinn unserer Existenz, dem Warum, Woher, Wozu« (Z. 40 f.) umfassend beantworten können.« [BY, 2019 (4)]

[Der dem Aufgabenset zugrunde liegende Text des Juristen HOST DREIER handelt von »Religion in der säkularen Moderne«. Das entnommene Zitat fungiert ausschließlich als »Stichwortgeber«. Inhaltlich hat die Aufgabe keine offensichtlichen Bezüge zum Text. Auch auf die »letzten Fragen« geht der Text nicht genauer ein als es in der Aufgabenstellung zitiert wird.]

Textaufgabe

»Setzen Sie sich mit Hubers Freiheitsverständnis auseinander.«⁵⁹ [NI, 2016, eA (1)]

Erweiterte Textaufgabe

»4. Setzen Sie sich unter Berücksichtigung ausgewählter Aspekte des Textes von Klaus-Peter Jörns mit dem Titelblatt des »SPIEGEL« (M2) auseinander.«⁶⁰ [NI, 2014, eA (1)]

Gestaltungsaufgabe

»In Ihrem Ort wird ein neues behindertengerechtes Sportzentrum eröffnet. Entwerfen Sie einen Beitrag für die Vereinszeitschrift des Sportvereins mit dem Titel »Schneller – besser – höher: aber natürlich!?, in dem Sie differenziert auf die Frage der Selbstoptimierung im Sport eingehen! Beziehen Sie dabei in kritischer Lektüre den Text (Seite 4 f.) sowie Gedanken eines christlichen Menschenbildes ein!« [BY, 2014 (2)]

B) Fragestellungen der Urteilsbildung

Mittels dieser Kategorie sollen die Fragestellungen der Urteilsbildung erfasst werden. Zum einen erfolgt dies über eine Zusammenstellung der Themen, zu denen eine Urteilsbildung stattfinden soll. Diese Themen werden durch eine qualitative Inhaltsanalyse rein induktiv erhoben. Überprüft wird dabei zum einen das Gesamtmaterial, um herauszufinden, ob sich klassische Themen oder Kanonbildungen der ethischen und religiösen Urteilsbildung im Religionsunterricht benennen lassen. Zum anderen werden die Bundesländer einzeln betrachtet, um herauszufinden, ob es bundeslandspezifische Schwerpunktsetzungen gibt.

Darüber hinaus sollen fünf induktiv entwickelte Merkmalsausprägungen erhoben werden, welche erfassen, ob die Fragestellung einen dezidierten Bezug

⁵⁹ Der Aufgabenstellung liegt ein Sachtext des Theologen WOLFGANG HUBER zugrunde.

⁶⁰ Der Aufgabenstellung liegt ein Text des evangelischen Theologen KLAUS-PETER JÖRNS zugrunde sowie das benannte Titelbild des SPIEGEL.

zur Religion aufweist. Zu prüfen ist, ob die jeweilige Fragestellung im Sinne der in Kapitel 2.1 vorgestellten Systematik als religiöse Fragestellung eingeordnet werden kann oder ob es sich um eine Fragestellung ohne religiösen Bezug handelt, die als ethisch-philosophische Fragestellung gefasst wird.⁶¹

Wesentlich für diese Kategorie ist, dass sie ausschließlich die Themen der Fragestellungen in den Blick nimmt und nicht die Reflexionsanforderungen, die in den Aufgabenstellungen oft enthalten sind, gerade wenn es sich um Fragestellungen ohne dezidiert religiösen Bezug handelt.

Die Unterscheidung zwischen Fragestellungen ohne dezidiert religiösen Bezug und religiösen Fragestellungen ist eine heuristische Unterscheidung, die vor allem deskriptiv einen bei der Analyse der Abituraufgabenstellungen zur Urteilskompetenz sichtbaren Unterschied zu erfassen sucht. Es bleibt dabei unbenommen, dass jede ethische Fragestellung durch ihre Perspektivierung zu einer christlichen Fragestellung werden kann (vgl. Kapitel 2.1). Das Unterscheidungskriterium für die Merkmalsausprägung ist jedoch, ob die zu verhandelnden Themen religiöser Art sind oder nicht. Neben diesen beiden Kategorien sind Fälle zu finden, bei denen eine Perspektive beurteilt werden soll, welche in der Regel in Form eines zugrunde liegenden Materials mit einer konkreten Person verknüpft ist, was in den Kategorien religiöse Perspektive auf eine Fragestellung ohne dezidiert religiösen Bezug, religiöse Perspektive auf eine religiöse Fragestellung sowie nicht-religiöse Perspektive auf eine religiöse Fragestellung erfasst wird. Die Aufforderung, dass eine nicht-religiöse Perspektive auf eine nicht dezidiert religiöse Fragestellung bewertet werden soll, kommt unter den Fällen der Studie nicht vor und wird daher auch nicht aufgeführt.

Kategorie B) Fragestellungen der Urteilsbildung

Kodierung	Merkmalsausprägung	Beschreibung
B1	Fragestellungen ohne dezidiert religiösen Bezug	Die Aufgabenstellung fordert eine Urteilsbildung zu einer Fragestellung, die nicht unmittelbar mit dem Thema Religion verknüpft ist. Eine Positionierung steht nicht zur Diskussion.
B2	religiöse Fragestellungen	Die Aufgabenstellung fordert eine Urteilsbildung zu einer Fragestellung, die mit dem Thema Religion unmittelbar verknüpft ist. Eine Positionierung steht nicht zur Diskussion.

⁶¹ Vgl. ERLER, Philosophie I.

Fortsetzung

Kategorie B) Fragestellungen der Urteilsbildung

Kodierung	Merkmalsausprägung	Beschreibung
B3	religiöse Perspektive auf eine Fragestellung ohne dezidiert religiösen Bezug	Gegenstand der Urteilsbildung ist eine Position zu einer Fragestellung ohne dezidiert religiösen Bezug. Die Position wird durch eine Person vertreten, welche sich – qua ihres Amtes oder durch Positionierungen im Text – als religiös definieren lässt.
B4	religiöse Perspektive auf eine religiöse Fragestellung	Gegenstand der Urteilsbildung ist eine Position zu einer religiösen Fragestellung. Die Position wird durch eine Person vertreten, welche sich – qua ihres Amtes oder durch Positionierungen im Text – als religiös definieren lässt.
B5	nicht-religiöse Perspektive auf eine religiöse Fragestellung	Gegenstand der Urteilsbildung ist eine Position zu einer religiösen Fragestellung. Die Position wird durch eine Person vertreten, welche sich nicht als religiös definieren lässt.

Ankerbeispielaufgaben Kategorie *Fragestellungen der Urteilsbildung*
Fragestellungen ohne dezidiert religiösen Bezug

»4 An Ihrer Schule wird die Ausstellung »Local heroes« gezeigt. Sie sind als Schülervertretung damit beauftragt, eine Rede bei der Eröffnung der Ausstellung zu halten. Entwerfen Sie eine Rede mit dem Titel »Tolle Typen, wozu? – Brauchen wir wirklich Vorbilder?«

Ziehen Sie dazu die Materialien 1–4 heran und berücksichtigen Sie dabei evangelische Einsichten zu Gott und Mensch!« [BY, 2014 (4)]

Religiöse Fragestellungen

»Beurteilen Sie den Wert von Gottesbeweisen für den Gottesglauben.« [TH, 2018, eA (2)]

Religiöse Perspektive auf eine Fragestellung ohne dezidiert religiösen Bezug

»3 Beurteilen Sie die Position Hubers zum Einsatz von Gewalt! Berücksichtigen Sie dabei Grundzüge biblischer Ethik sowie Denkmodelle zum politischen Engagement von Christen!« [BY, 2016 (4)]

Religiöse Perspektive auf eine religiöse Fragestellung

»4. Setzen Sie sich mit der Position Steffenskys auseinander.«⁶² [NI, 2017, eA (2)]

Nicht-religiöse Perspektive auf eine religiöse Fragestellung

»4. Setzen Sie sich in einem Brief an Abdel-Samad damit auseinander, in welcher Form uns insbesondere in der Folge der Aufklärung die Religionen in die Zukunft begleiten können.« [NI, 2018, eA (2)]

C) Reflexionskriterien

In den EPA wird im AFB III die Anforderung formuliert, die »eigene[] Urteilsbildung unter Beachtung biblischer, theologischer und ethischer Kategorien« zu reflektieren.⁶³ Die Analysekategorie *Reflexionskriterien* prüft daher, inwiefern die untersuchten Aufgabenstellungen spezifische Reflexionsanforderungen stellen.

Im Datenmaterial wurden verschiedene Formen gefunden, die zu einer Reflexion der eigenen Urteilsbildung auffordern. Es wird definiert, dass folgende Fälle als Reflexionsanforderungen betrachtet werden:

- 1.) die Aufforderung, etwas in die Urteilsbildung einzubeziehen – seien es christliche oder philosophische Positionen oder Konzepte, Materialien, persönliche Erfahrungen etc.

Folgende Formulierungen illustrieren dies beispielhaft: »[...] Beziehen Sie dabei grundlegende Aspekte philosophischer und christlicher Ethik mit ein!« [BY, 2017 (1)] oder »[...] Beziehen Sie dabei in kritischer Lektüre den Text [...] ein!« [BY, 2014 (2)]. Abzugrenzen ist jedoch der Fall, dass über die gegebene Position selbst geurteilt werden soll. Dann ist diese nicht mehr Reflexionskriterium, sondern Urteilsgegenstand (vgl. Kategorie B): In der Aufgabenstellung, »Setzen Sie sich mit der Position des Leserbriefes (M 2) auseinander« [NI, 2018, gA (1)], ist demnach keine Reflexionsaufforderung enthalten, sondern die gegebene Position ist der Urteilsgegenstand. Als Unterscheidungsmerkmal dient die Formulierung »beziehen Sie ... ein«. Sobald diese vorhanden ist, liegt eine Reflexionsaufforderung vor, da das Einbeziehen der Materialien einen Anstoß gibt, die eigene Urteilsbildung in Beziehung zu der anderen Perspektive zu setzen – folglich zu reflektieren. Ebenso müssen Materialien bei ihrem Einbezug ins Verhältnis zur eigenen Meinung gesetzt werden, wodurch eine Reflexion entsteht.

- 2.) konkrete Kontexteinbettungen der Aufgaben, die einen spezifischen Blick auf die Fragestellung mit sich bringen

Ein Beispiel bietet folgende Aufforderung:

⁶² Zugrunde liegt ein Text FULBERT STEFFENSKYS, Professor für Religionspädagogik, der sich mit dem Verhältnis von Sünde und Vergebung sowie Gott und Mensch auseinandersetzt.

⁶³ Vgl. EPA, 12.

»[...] Entwerfen Sie für das Fach Evangelische Religionslehre einen Beitrag, der sich mit Chancen und Grenzen des technisch Machbaren auseinandersetzt! [...]« [BY, 2018 (4)].

Die spezifische Kontexteinbettung gibt vor, dass der Prüfling seine Urteilsbildung im Kontext dieser Sichtweise – im vorliegenden Beispiel im Kontext des Faches, für das er den Artikel schreibt – reflektieren muss.

Auf Grundlage dieser Definition konnten induktiv folgende Merkmalsausprägungen im Datenmaterial gefunden werden:

Kategorie C: Reflexionsanforderungen		
Kodierung	Merkmalsausprägung	Beschreibung
C0	Keine expliziten Reflexionsanforderungen	Die Aufgabenstellung fordert keine zusätzlichen Reflexionsanforderungen.
C1	Christliche Sichtweisen als Reflexionsanforderung	Die Aufgabenstellung fordert den Einbezug christlicher Sichtweisen bzw. geeigneter Bibelstellen oder christlicher Anthropologie, ohne dass diese selbst Urteilsgegenstand ist.
C2	Nicht-christliche Sichtweisen als Reflexionsanforderung	Die Aufgabenstellung fordert den Einbezug nicht christlicher Sichtweisen, ohne dass die diese selbst Urteilsgegenstand ist.
C3	Materialien als Reflexionsanforderung	Der Aufgabenstellung liegen Materialien zugrunde, die der Reflexion der Urteilsbildung dienen. Es soll nicht über das Material selbst geurteilt werden.
C4	Christliche Kontextualisierung der Aufgabenstellung als Reflexionsanforderung ⁶⁴	Die Aufgabenstellung gibt eine christliche Kontextualisierung vor, welche zu einer Reflexion der eigenen Urteilsbildung auffordert.
C5	Persönliche Erfahrungen	Die Aufgabenstellung fordert den Einbezug persönlicher Erfahrungen in den Urteilsbildungsprozess.

⁶⁴ Eine Kontextualisierung anderer Art mit bspw. biologischem, medizinischem oder philosophischem Fokus kommt im Datenmaterial nicht vor, so dass diese Merkmalsausprägung auf eine christliche Kontextualisierung beschränkt werden konnte.

Zur Kodierung werden ausschließlich die Formulierungen der Aufgabenstellungen betrachtet. Das Material wird an dieser Stelle nicht hinsichtlich seiner Ausrichtung (christlich oder nicht-christlich) analysiert (vgl. dazu Kapitel 6.2.4). Das Vorkommen mehrerer Reflexionsanforderungen in einer Aufgabenstellung ist möglich, so dass es in den Merkmalsausprägungen C1-C5 zu Mehrfach-Zuordnungen kommen kann. Durch die Kategorie C0 kann man trotz dieser Mehrfach-Zuordnungen errechnen, auf wie viele Aufgaben sich die Reflexionsanforderungen verteilen. Die Aufforderung, Beispiele anzuführen, wird als Stützung der Argumente und nicht als Reflexionsaufforderung gewertet.

Ankerbeispiele Kategorie *Reflexionsanforderungen*

Keine expliziten Reflexionsanforderungen

»4. Entwerfen Sie einen Kommentar zu Käbmans friedensethischer Position.«⁶⁵ [NI, 2019, eA (1)]

Christliche Sichtweisen als Reflexionsanforderungen (C1)/Materialien als Reflexionsanforderungen (C3)/Persönliche Erfahrungen als Reflexionsanforderung(C5)

»Entwerfen Sie in Auseinandersetzung mit den Worten von Margot Käbmann eine Rede zur Abschlussfeier der Absolventen Ihres Jahrgangs.

Beziehen Sie die Materialien, geeignete Texte aus der Bibel sowie Ihre persönlichen Erfahrungen in Ihre Festrede mit ein.«⁶⁶ [TH, 2014, eA (1)].

Nicht-christliche Sichtweisen als Reflexionsanforderungen

»3. Die Religionskritik weist zu Recht auf das in vieler Hinsicht katastrophale Versagen der Religionen hin. Aber Religionskritik kann Religion nicht ersetzen [...].« (Z. 19–20)

Nehmen Sie zu dieser Aussage vor dem Hintergrund eines religionskritischen Konzeptes Stellung.« [NI, 2016, gA (2)]

Christliche Kontextualisierung der Aufgabenstellung als Reflexionsmedium

»3. Nehmen Sie in Form eines Beitrags für die Internetseite »evangelisch.de« Stellung zu Hintzes Position.« [NI, 2014, gA (1)]

D) Textsorten der zugrunde liegenden Materialien

Die den Aufgaben zur Urteilsfähigkeit zugrunde liegenden Materialien stammen aus unterschiedlichen Textsorten, die von Sachtexten bis zu Gedichten reichen. Kategorie D untersucht die den Aufgabenstellungen zugrunde liegenden Text-

⁶⁵ Der Aufgabenstellung liegt ein Kommentar der ehemaligen Ratsvorsitzenden der EKD, MARGOT KÄBMANN, zugrunde.

⁶⁶ Der Aufgabe liegt ein Zitat MARGOT KÄBMANNs über die politische Dimension des Christentums zugrunde, in dem sie sich als »Weltverbesserin« bekennt.

sorte, um Unterschiede zwischen den drei untersuchten Bundesländern beschreiben zu können. Induktiv konnten folgende Merkmalsausprägungen im Material gefunden werden:

Kategorie D: Textsorten der zugrunde liegenden Materialien		
Kodierung	Merkmalsausprägung	Beschreibung
D0	Kein Materialbezug	Der geforderten Urteilsbildung liegt kein Materialbezug zugrunde. (Die Kategorie entspricht damit der Kategorie A1.)
D1	informierender Texte	Der der Aufgabenstellung zugrunde liegende Text besitzt informierenden Charakter und ist in der Regel in einer Zeitung erschienen. Folgende Textsorten werden unter dieser Kategorie gefasst: – Bericht – Nachricht – Reportage – Begleittext zu einer Ausstellung – Einführung in Rundfunksendung
D2	meinungsäußernde Texte	Der der Aufgabenstellung zugrunde liegende Text ist deutlich subjektiv gefärbt und meinungsäußernd. Die Texte sind in der Regel nicht von einem Verlag publiziert oder aber deutlich als Meinung gekennzeichnet, wie es beim Leserkommentar der Fall ist. Folgende Textsorten werden unter dieser Kategorie gefasst: – Forenbeitrag im Internet – Leserkommentar – Leserfrage mit Antwort

Fortsetzung

Kategorie D: Textsorten der zugrunde liegenden Materialien

Kodierung	Merkmalsausprägung	Beschreibung
D3	Ethisch oder theologischer Sachtext	Der der Aufgabenstellung zugrunde liegende Text ist ein theologischer Sachtext. Religiöse Fragestellungen werden in diesem auf sachlicher Ebene unter Bezugnahme auf Argumente verhandelt. Folgende Textsorten werden unter dieser Kategorie gefasst, sofern sie das Kriterium einer argumentativen Auseinandersetzung mit einer ethischen oder religiösen Fragestellung erfüllen: – Textauszüge aus Monografien – Essays – Festreden – Interviews – Predigtauszüge – (Fach-)Zeitschriftenbeiträge – Kommentare – Stellungnahmen der EKD – Aufruf der Arbeitsgemeinschaft christlicher Kirchen in Bayern zur Hilfe für Geflüchtete
D4	Lyrische Texte	Der der Aufgabenstellung zugrunde liegende Text ist lyrischer Art. Folgende Textsorten werden unter dieser Kategorie gefasst: – Gedichte – Liedtexte – Glaubenszeugnisse/Bekenntnisse einzelner Personen⁶⁷ – Gebete
D5	Zitat	Der Aufgabenstellung liegt kein Textauszug, sondern lediglich ein Zitat zugrunde, welches aus einem oder zwei Sätzen besteht und nicht aus der Bibel stammt.

⁶⁷ Die Gründe für diese Zuordnung sind die in den Glaubensbekenntnissen verstärkt auftretenden rhetorischen Mittel und sprachlichen Bilder. Das Apostolische Glaubensbekenntnis wurde hingegen der Kategorie D6 »Heilige Texte« zugeordnet.

Fortsetzung

Kategorie D: Textsorten der zugrunde liegenden Materialien

Kodierung	Merkmalsausprägung	Beschreibung
D6	Heilige Texte	Der Aufgabenstellung liegt eine biblische Perikope zugrunde. ⁶⁸ Darüber hinaus wurde das Apostolische Glaubensbekenntnis dieser Kategorie zugeordnet.
D7	Bild	Der Aufgabenstellung liegt ein Bild zugrunde.
D8	Statistik	Der Aufgabenstellung liegt eine Statistik zugrunde.

Da sich mitunter eine Aufgabenstellung auf mehrere Materialien bezieht, schließen sich die Merkmalsausprägungen D1–D7 gegenseitig nicht aus. Ebenso kann die gleiche Textsorte einer Aufgabe mehrfach zugrunde liegen. Die Kategorie bezieht sich ausschließlich auf die Materialien und nicht auf in der Aufgabenstellung enthaltene Zitate.

5.3.3 Kategorien zur Analyse der Erwartungshorizonte

E) Passung zwischen geforderter und erwarteter Urteilskompetenz

Kategorie E erhebt, inwieweit die in den Aufgabenstellungen geforderte Anforderung einer Urteilsbildung in den Erwartungshorizonten wieder aufgegriffen wird. Dazu werden die Erwartungshorizonte aller 290 im Untersuchungszeitraum gestellten Aufgabenstellungen hinsichtlich ihrer Forderungen einer persönlichen Urteilsbildung der Prüflinge untersucht und mit den Anforderungen der Aufgabenstellungen abgeglichen. Ziel ist es, die Aufgaben zu erfassen, welche zwar eine Urteilsbildung in ethischen und religiösen Fragestellungen fordern, sie aber in den Erwartungshorizonten nicht einholen bzw. die Aufgabenstellungen, deren Erwartungshorizonte eine persönliche Urteilsbildung vorgeben, die aus der Formulierung der Aufgabenstellungen nicht hervorgeht. Mögliche Gründe für eine fehlende Passung werden an den konkreten Fällen diskutiert.

⁶⁸ Die Aufforderung, geeignete Bibelstellen mit einzubeziehen, ohne dass diese angegeben sind, wird nicht als Material gewertet, sondern ist unter Kategorie C Reflexionskriterien erfasst.

Kategorie E: Passung zwischen geforderter und erwarteter Urteilsfähigkeit

Kodierung	Merkmalsausprägung	Beschreibung
E1	Passung gegeben	Die Aufgabenstellung fordert eine Urteilsbildung in Bezug auf eine ethische oder religiöse Fragestellung, die im Erwartungshorizont auch eingeholt wird.
E2	Aufgabenstellung nicht passgenau	Aufgabenstellung fordert keine eigenständige Urteilsbildung, während der Erwartungshorizont diese verlangt.
E3	Erwartungshorizont nicht passgenau	Der Erwartungshorizont erwähnt keine eigenständige Urteilsbildung in Bezug auf religiöse und ethische Fragestellungen, obwohl die Aufgabenstellung diese fordert.

Auf die Formulierung von Ankerbeispielen wird verzichtet, da die Darstellung der Erwartungshorizonte umfangreich ist und die Fälle fehlender Passgenauigkeit ausführlich in den Ergebnissen diskutiert werden (vgl. Kapitel 6.3.1).

F) Qualitätsmerkmale von Urteilsbildung in den Erwartungshorizonten

Kategorie F erfasst induktiv, welche Qualitätsmerkmale für Urteilsbildung in kontroversen ethischen oder religiösen Fragestellungen in den Erwartungshorizonten genannt werden. Als Qualitätsmerkmale werden dabei zum einen jene Angaben erfasst, die Teilaspekte benennen, welche im Urteilsentscheid enthalten sein sollen, wie die Stellungnahme selbst, eine Begründung oder ein Aufgabenbezug. Zum anderen werden Textstellen kategorisiert, welche über die Qualität dieser Teilaspekte Aussagen treffen, wie beispielsweise »differenziert«, »schlüssig« oder »plausibel«. Die Kategorisierung unterscheidet zwischen aufgabentextunspezifischen Qualitätsmerkmalen, aufgabentextspezifischen Qualitätsmerkmalen und explizit im Erwartungshorizont nivellierte Qualitätsmerkmale:

Kategorie F: Qualitätsmerkmale von Urteilsbildung

Kodierung	Merkmalsausprägung	Beschreibung
F1	Aufgabentextunspezifische Qualitätsmerkmale	Die im Erwartungshorizont benannten Qualitätsmerkmale zur Urteilsbildung in kontroversen ethischen und religiösen Fragestellungen sind nicht explizit in der Aufgabenstellung gefordert.
F2	Aufgabentextspezifische Qualitätsmerkmale	Die im Erwartungshorizont benannten Qualitätsmerkmale zur Urteilsbildung in kontroversen ethischen und religiösen Fragestellungen wurden explizit in der Aufgabenstellung gefordert.
F3	Explizit nivellierte Qualitätsmerkmale	Im Erwartungshorizont werden Merkmale explizit nivelliert.

Ankerbeispiel Kategorie *Qualitätsmerkmale von Urteilsbildung*

»3. Entwerfen Sie einen Brief an Nikolaus Schneider zu dessen Sicht des Menschen.«⁶⁹ [NI, 2016, gA (1)]

»Diese Teilaufgabe ist schwerpunktmäßig dem AFB III zuzuordnen.

Zum einen erfordert die Aufgabenstellung die Beachtung des formalen Rahmens eines Briefes; zum anderen sollen sich die Prüflinge kreativ mit der auf den Text bezogenen Aufgabenstellung auseinandersetzen.

Die Aufgabe überlässt den Prüflingen u. a. die Entscheidung,

- sich gegen die Sicht Schneiders zu positionieren und z. B. einen anderen anthropologischen Entwurf darzustellen (z. B. Verantwortung und Freiheit des Menschen auch ohne Gottesglauben),
- andere biblisch-christliche Akzente des Menschseins herauszustellen bzw. anders zu gewichten, als Schneider dies tut, z. B. Betonung der Gnade Gottes gegenüber dem Menschen oder die Rechtfertigung durch den Glauben,
- einzelne Aspekte des Textes besonders in den Fokus zu nehmen, wie z. B. den, dass dem Menschen das für ein gelingendes Leben Entscheidende von Gott gesagt werde, den der Gefahr der Selbstüberschätzung oder den der häufig schwierigen Entscheidungsfindung angesichts der Probleme in der Welt, [sic!]

⁶⁹ Der Aufgabenstellung liegt ein Text des evangelischen Theologen NIKOLAUS SCHNEIDER zugrunde.

Im Brief sollten der Bezug zum Text Schneiders und eine eigene Position des Prüflings deutlich werden.

[...]

Eine Leistung kann mit »gut« bewertet werden, wenn [...]

- der Brief überzeugend und einfallsreich gestaltet ist und eine ausführliche, aspektreiche und nachvollziehbare Bezugnahme zu Schneiders Menschenbild darstellt.

[...]« [E-NI, 2016, gA (1)]

Aufgabentextunspezifische Qualitätsmerkmale:

- Positionierung
- Eigenständigkeit
- Ausführlichkeit
- (inhaltlich) aspektreich
- Nachvollziehbarkeit

Aufgabentextspezifische Qualitätsmerkmale:

- Briefform
- Kreativität
- überzeugende und einfallsreiche Gestaltung des Briefes
- Materialbezug
- Aufgabenbezug,

Explizit nivellierte Qualitätsmerkmale:

- im vorliegenden Beispiel keine

G) Explizite Hinweise zum Umgang mit Einstellungen

Kategorie G analysiert die Erwartungshorizonte der Urteilsbildungsaufgabenstellungen danach, ob sie explizite Hinweise zum Umgang mit den Einstellungen und Positionierungen der Prüflinge enthalten. Die Kategorie nimmt ausschließlich Textstellen auf, die die Anforderungen der Aufgabe reflektieren, also gewissermaßen »Metatext« darstellen. Reine inhaltliche Abgrenzungen zwischen Pro- und Kontraargumentationen – die ebenso die Offenheit des Urteilsentscheides signalisieren – werden in dieser Kategorie nicht betrachtet, sondern sind in der Kategorie *Formulierung konträrer Positionen in den Erwartungshorizonten* zu finden.

Kategorie G: Qualitätsmerkmale von Urteilsbildung

Kodierung	Merkmalsausprägung	Beschreibung
G0	Keine expliziten Hinweise	Es werden im Erwartungshorizont der Aufgabenstellungen, die eine ethische oder religiöse Urteilsbildung fordern, keine expliziten Hinweise zum Umgang mit Einstellungen oder Positionierungen der Prüflinge gegeben.
G1	Explizite Hinweise	Der Erwartungshorizont einer Aufgabenstellung zur Urteilsfähigkeit in ethischen oder religiösen Fragestellungen enthält explizite Hinweise zum Umgang mit Einstellungen oder Positionierungen der Prüflinge bei der Korrektur und/oder Bewertung.

Ankerbeispiele Kategorie *Explizite Hinweise zum Umgang mit Einstellungen*
Keine expliziten Hinweise

Der Erwartungshorizont enthält keine expliziten Hinweise.

Explizite Hinweise

»Aufgrund der eigenen Position des Prüflings und des vorangegangenen Unterrichts sind verschiedene Bearbeitungen dieser Aufgabe möglich« [E-NI, 2015, gA (2)].

H) Formulierung konträrer Positionen in den Erwartungshorizonten

Kategorie H fragt, ob die Erwartungshorizonte der Aufgaben zur Urteilsfähigkeit zwei oder mehrere konträre Positionen zur Fragestellung inhaltlich explizit formulieren oder ob ausschließlich allgemeine Qualitätskriterien genannt werden. Die Kategorie knüpft damit eng an die Fragestellung an, wie mit der Offenheit zur Urteilsbildung auffordernder Aufgabenstellungen umgegangen wird.

Kategorie H: Formulierung konträrer Positionen in den Erwartungshorizonten

Kodierung	Merkmalsausprägung	Beschreibung
H0	Konträre Positionen nicht formuliert	Im Erwartungshorizont der Aufgabenstellung werden die möglichen konträren Positionen zur Fragestellung nicht formuliert.
H1	Konträre Positionen formuliert	Im Erwartungshorizont der Aufgabenstellung werden mindestens zwei mögliche konträre Positionen zur Fragestellung formuliert.

Ankerbeispiele Kategorie *Formulierung konträrere Positionen in den Erwartungshorizonten*:

Konträre Positionen nicht formuliert

Erwartungshorizont der Aufgabe: »Lässt sich mit der Bergpredigt Politik gestalten?«

»**Erörtern:** zu einer vorgegebenen Problemstellung eine reflektierte Auseinandersetzung führen und zu einer abschließenden, begründeten Bewertung gelangen

- schlüssige und verständliche Gedankenentwicklung
- Bezugnahme auf die Aussagen der Bergpredigt
- Bezugnahme auf selbst gewählte Beispiele
- Argumente ausführlich entfalten
- sachlich, klar und präzise formulieren
- kohärenter Aufbau« [E-TH, 2017 (2)]

Konträre Positionen formuliert

Erwartungshorizont zur Aufforderung, die friedensethische Position MARGOT KÄBMANNs zu bewerten.

»Möglich ist, dass der Prüfling in dem Kommentar der Haltung Käbmans, ihren Gedanken und Forderungen zustimmt. Zur Unterstützung von Käbmans friedensethischer Position kann der Prüfling folgende Aspekte anführen:

- Das Gebot der Feindesliebe ist eine zentrale Forderung Jesu in der Bergpredigt, die grundsätzlich gilt.
- Nachfolge Jesu beinhaltet auch den Einsatz für ein friedliches Zusammenleben. [...]

Demgegenüber ist eine kritische Würdigung ihrer Position oder einzelner Aspekte ihrer Predigt denkbar. Mögliche Einwände können dabei sein:

- Das Gebot der Feindesliebe ist eine utopische Forderung und überfordert menschliches Handeln.
- Eine militärische Intervention erweist sich in der Auseinandersetzung mit dem Terrorismus in der Realität als angemessener (Hinweis auf die Lehre vom »gerechten Krieg«). [...]« [E-NI 2019 gA (1)]

5.4 Analysevorgehen und Methodenreflexion

Nachdem das Kategoriensystem nun inhaltlich vorgestellt wurde, soll seine Entstehung, das Analysevorgehen sowie der Ertrag des Kategoriensystems reflektiert werden, bevor die Ergebnisse der Studie dargestellt werden.

5.4.1 Beantragung der Einsicht in die Abituraufgabenstellungen

Die Einsichtnahme in die Abituraufgabenstellungen und ihre Erwartungshorizonte bei den Kultusministerien der Länder Bayern, Niedersachsen und Thüringen erfolgte in zwei Phasen: Zu Beginn des Arbeitsprozesses 2016 wurde die Einsicht in die Jahre 2014 bis 2016 beantragt und von allen drei Bundesländern bewilligt.⁷⁰ Anhand dieser Aufgabenstellungen wurden die deduktiv entwickelten Hauptkategorien induktiv durch Subkategorien ergänzt und weiterentwickelt. Beispiele für diesen Prozess werden im Folgenden gegeben (vgl. insbesondere Kapitel 5.4.3).

Um möglichst aktuelle Aussagen über die Prüfpraxis und die Bewertungskriterien von Urteilsfähigkeit treffen zu können, wurde 2019 nochmals die Einsicht in die Abituraufgaben und Erwartungshorizonte der Jahre 2017–2019 erfolgreich angefordert. Die neu hinzugekommen Aufgaben haben dazu geführt, dass alle Kategorien nochmals geprüft und Analysevorgänge wiederholt durchgeführt wurden, was zu einer höheren Validität der Analyse beiträgt. Induktive Anpassungen des Kategoriensystems wurden auch in diesem Prozess vorgenommen, so dass das Kategoriensystem durchweg flexibel gehalten wurde.

5.4.2 Auswahl der Aufgabenstellungen zur Urteilsbildung

Das Interesse der Studie, die Prüfungskonventionen sowie Bewertungskriterien von Urteilsbildung in Bezug auf kontroverse ethische und religiöse Fragestellungen zu untersuchen, ist deduktiv, also unabhängig vom Datenmaterial, entstanden (zur Herleitung der Fragestellung vgl. Kapitel 2). Es galt folglich zunächst, all jene im Untersuchungszeitraum in den drei Bundesländern gestellten Aufgabenstellungen zu benennen, welche eine ethische und/oder religiöse Urteilsbildung fordern. Diese bilden die Fälle der Studie (vgl. Kapitel 5.2.1).⁷¹ Zu diesem Zweck wurde zunächst eine Systematisierung aller Aufgabenstellungen vorgenommen, indem sie tabellarisch den drei AFB zugeordnet wurden. Diese Zuordnung erfolgte anhand der Operatoren der Aufgabenstellungen. Durch die in den EPA festgeschriebene Verortung von Urteilsfähigkeit in den AFB III (vgl. Kapitel 3.2.2) hätte man Urteilsbildungsaufgaben ausschließlich in der Spalte zu AFB III vermuten können. Der zunächst gewählte Weg, die Aufgabenstellung zur Urteilsbildung in kontroversen ethischen und religiösen Fragestellungen anhand

⁷⁰ Für die Übermittlung der Aufgabenstellungen und der dazugehörigen Erwartungshorizonte sei an dieser Stelle nochmals ausdrücklich dem THÜRINGER MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT, dem BAYERISCHEN STAATSMINISTERIUM FÜR UNTERRICHT UND KULTUS sowie dem NIEDERSÄCHSISCHEN KULTUSMINISTERIUM gedankt.

⁷¹ Zum Begriff des Falls vgl. KUCKARTZ, Qualitative Inhaltsanalyse, 49.

der Operatoren zu bestimmen, die laut den EPA eine Urteilsbildung fordern (»sich auseinandersetzen mit«, »beurteilen«, »bewerten« etc.; vgl. Kapitel 5.2.2), erwies sich aufgrund von vorhandenen Unstimmigkeiten zwischen den verwendeten Operatoren und den inhaltlichen Anforderungen der Aufgaben als nicht gangbar: Die Analyse der Aufgabenstellungen zeigte, dass sich auch in Aufgabenstellungen mit anderen Operatoren Aufgaben finden lassen, welche zur ethischen und religiösen Urteilsbildung auffordern und dass diese über alle drei AFB hinweg verteilt sind. Es bedurfte folglich eines neuen Auswahlkriteriums, welches in einer inhaltlichen Definition gefunden wurde (vgl. Kapitel 5.2.2). Um die festgestellten Unstimmigkeiten zwischen den Operatoren und den inhaltlichen Aufforderungen der Aufgabenstellung fassen zu können und um sie mit den Anforderungen der Erwartungshorizonte abzugleichen, wurde die Analyse-kategorie *Passung zwischen geforderter und erwarteter Urteilsfähigkeit* erstellt.

Die Prüfung aller Aufgabenstellungen anhand der inhaltlichen Definition führte in einigen Fällen zu Uneindeutigkeiten, welche teilweise durch eine Ausschärfung der Definition behoben werden konnten. Die Aufgabenstellungen, in denen dies nicht möglich war, wurden erfasst und mit einer Lehrkraft diskutiert, um bei der Zuordnung der Gefahr zu entgehen, das Alltagsverständnis von Urteilsfähigkeit zu unterlaufen. Die Entscheidung einer Aufnahme oder eines Ausschlusses wurden in Kapitel 5.2.3 »Grenz- und Sonderfälle« begründet. Nach mehrmaligen Analysedurchgängen aller Aufgabenstellung wurden 42 Aufgabenstellungen als Urteilsbildungsaufgaben definiert. Der Studie liegen folglich 42 zu analysierende Fälle zugrunde.

Die für die Auswahl der Aufgabenstellung zu ethischer und religiöser Urteilsfähigkeit erstellte tabellarische Zuordnung zu den drei AFB diente im weiteren Analyseverfahren zudem der Erhebung quantitativer Aspekte, da sie alle 290 im Untersuchungszeitraum gestellten Aufgaben erfasst (vgl. Kapitel 6.1).

5.4.3 Erstellung des Kategoriensystems und Auswertung der Daten

An der folgenden Darstellung soll deutlich werden, dass die Entwicklung des Kategoriensystems einen *zirkulären Prozess* darstellte, welcher das Ziel verfolgte, das Analysesystem so zu konzipieren, dass es den Untersuchungsgegenstand möglichst treffend zu erfassen und zu beschreiben vermag. Der Ablauf des Analyseprozesses orientierte sich an den von UDO KUCKARTZ definierten sieben Phasen einer inhaltlich strukturierenden qualitativen Inhaltsanalyse, wie sie in der folgenden Grafik dargestellt sind.

Nach der Sichtung des Materials wurden die Hauptkategorien des Kategoriensystems erstellt, welche sich auf die Aufgabenstellungen beziehen. Es handelt sich um deduktiv entstandene thematische Kategorien. Kategorie A *Aufgabenarten* ist an die EPA-Vorgaben angelehnt und wurde nicht weiter induktiv



Abb. 10: »Ablaufschema einer inhaltlich strukturierenden Inhaltsanalyse«. Übernommen aus: KUCKARTZ, Qualitative Inhaltsanalyse, 100.

ergänzt. Zur Erfassung der Themen der Urteilsbildung (Kategorie B) wurden diese zunächst anhand von Stichpunkten erfasst und daraufhin analysiert, ob diese thematisch oder aufgrund der Perspektive einen Bezug zu Religion aufweisen oder nicht.

Eine Kategorie, die danach fragte, ob eine christliche Weltsicht in den Aufgabenstellungen eine Rolle spielen, musste konkretisiert werden, da alle im Untersuchungszeitraum gestellten Aufgabenstellungen zu ethischer oder religiöser Urteilsbildung einen religiösen Bezug aufweisen. Mit Blick auf das Datenmaterial wurde in diesem Zusammenhang die Bedeutung von Reflexionsanforderungen in den Aufgabenstellungen sichtbar. Es wurde daher eine Kategorie entwickelt, welche die Reflexionsanforderungen der Aufgabenstellungen abbildet (vgl. Kategorie C). Durch diese Kategorie, deren Merkmalsausprägungen induktiv entstanden sind, konnte auch die Rolle christlicher Weltsichten in den Aufgaben über die Thematik der Fragestellung hinaus erfasst werden.

Die Textsorten der Materialien (Kategorie D) wurden zunächst in Stichpunkten notiert und im Anschluss induktiv zu Gruppen zusammengefasst, welche die Merkmalsausprägungen der Kategorie bilden.

Die erste Kategorie, welche sich auf die Erwartungshorizonte bezieht (Kategorie E), wurde induktiv aus den beschriebenen Unstimmigkeiten in der Nutzung der Operatoren erstellt (vgl. Kapitel 5.4.2). Sie bezieht sich auf die Passung zwischen Aufgabenstellung und Erwartungshorizont im Hinblick auf eine Urteilsbildung. Zur Analyse dieser Kategorie wurden die Erwartungshori-

zonte aller 290 Aufgabenstellungen der Auswahlinheit durchgesehen und auf ihre Forderungen zu Urteilsfähigkeit hin geprüft.

Eine Unterteilung in »von der Aufgabenstellung geforderte Qualitätsmerkmale« und »aufgabentextunspezifische Qualitätsmerkmale« der Kategorie F ergab sich deduktiv aus dem Interesse der Fragestellung heraus, während die beiden Kategorien G und H zum expliziten Umgang mit Einstellungen und Haltungen sowie zur Formulierung konträrer Positionen induktiv entstanden sind: Da entsprechende Fälle im Datenmaterial gefunden wurden, sollte erhoben werden, wie oft diese vorkommen und in welcher Form.

Die grundlegenden Strukturierungsdimensionen des Materials bildeten die im Untersuchungszeitraum gestellten Aufgaben zur Urteilsfähigkeit (die Fälle) auf der einen und die Kategorien auf der anderen Seite. Es wurden angelehnt an KUCKARTZ zur Analyse der Daten zwei Themenmatrizen erstellt (eine in Bezug auf die Aufgabenstellungen und eine in Bezug auf die Erwartungshorizonte), in denen jeweils die Fälle in der Spalte und die Kategorien in der Zeile stehen.⁷² Die Tabellen erfassen die jeweiligen Merkmalsausprägungen des Falls. Diese wurde zum einen über die Kürzel der Subkategorien (zu Kategorie A: A0, A1, A2...) zum anderen inhaltlich über Textstellen erfasst.

Die Auswertung der Themenmatrizen wurde zunächst ausschließlich aus einer vertikalen Perspektive vorgenommen: Sie folgte einer kategorienorientierten Analyse, welche die Spalten im Blick hatte und zur jeweiligen Kategorie themenorientiert Aussagen über alle Fälle der Studie generierte. Die Ergebnisse dieser Analyse der einzelnen Kategorien werden in Kapitel 6 dargestellt.⁷³ Eine wesentliche Rolle spielten dabei quantifizierende Materialanalysen, die Aussagen über die Häufigkeiten verschiedener Phänomene zulassen.⁷⁴ Quantifizierende Auswertungen können im Kontext qualitativer Forschungen aufgrund der geringen Fallzahlen und aufgrund komplexer, schwer zählbarer Textstellen kritisch hinterfragt werden.⁷⁵ Daher ist es für die qualitative Forschung von Bedeutung, die Analyse weiter zu vertiefen. Dies wurde auch in der Reflexion der Ergebnisse dieser Studie deutlich: Der Kontext, in welchen die einzelnen Aufgabenstellungen im Aufgabenset eingebettet sind, sowie das Zusammenspiel einzelner Aspekte der Erwartungshorizonte konnten durch eine vertikale Perspektive nicht hinreichend erfasst werden. Die kategorienorientierte Analyse

⁷² Vgl. KUCKARTZ, Qualitative Inhaltsanalyse, 49.

⁷³ Vgl. zu vertikalen und horizontalen Perspektiven im Auswertungsprozess qualitativer Studien a. a. O., 50.

⁷⁴ Vgl. zu quantifizierenden Materialübersichten auch CHRISTIANE SCHMIDT, Analyse von Leitfadeninterviews, in: BARBARA FRIEBERTSHÄUSER/ANTJA LANGER/ANNEDORE PRENGEL (Hrsg.), Handbuch. Qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim 42013, 473–486, hier 481 f.

⁷⁵ Vgl. ebd.

wurde daraufhin durch eine fallorientierte Analyse ergänzt, welche durch eine horizontale Perspektive auf exemplarische Fälle nochmals kontextualisierte Einsichten liefert und in der Lage ist, Zusammenhänge zu problematisieren (vgl. Kapitel 7).⁷⁶ Dabei werden alle Kategorien zu einem Fall betrachtet. CHRISTIANE SCHMIDT spricht von »vertiefenden Einzelfallinterpretationen«, welche besonders interessante Einzelfälle zur Illustration der verallgemeinernden Befunde mit dem Ziel untersucht, Zusammenhänge darzustellen und theoretische Schlussfolgerungen zu ziehen.⁷⁷ Die vertiefenden Fallinterpretationen erfolgen nach SCHMIDT unter einer bestimmten, ausgewählten Fragestellung, welche sich im Kontext dieser Studie auf die in den Erwartungshorizonten genannten Qualitätsmerkmale für Urteilsfähigkeit bezieht und nach deren Maßstabsgehalt fragt (vgl. Kapitel 7).⁷⁸ Die Ergebnisse sowohl der kategorienorientierten Analyse als auch der fallbasierten Analyse werden in den folgenden zwei Kapiteln präsentiert.

⁷⁶ Vgl. KUCKARTZ, Qualitative Inhaltsanalyse, 115–117.

⁷⁷ Vgl. SCHMIDT, Analyse von Leitfadeninterviews, 482–484.

⁷⁸ Vgl. a. a. O., 482.

6 Kategorienorientierte Ergebnisse

Das vorliegende Kapitel dient dazu – wie bereits im Vorangegangenen beschrieben – die Ergebnisse einzelner Kategorien darzustellen. Die Darstellung ist weitestgehend deskriptiv und sieht von einer Interpretation zunächst ab. Dazu werden die Ergebnisse qualitativer Aspekte (Kapitel 6.1), die Ergebnisse, die sich auf die Aufgabenstellungen beziehen (Kapitel 6.2) und die Ergebnisse der Analysekatoren der Erwartungshorizonte (Kapitel 6.3) nacheinander vorgestellt. Es erfolgt zunächst jeweils eine Gesamtdarstellung, welche dann in Hinblick auf die drei untersuchten Bundesländer ausdifferenziert wird.

Interpretierende Perspektiven werden in den drei darauffolgenden Kapiteln aufgegriffen: In Kapitel sieben, indem einzelne Fälle betrachtet und diskutiert werden, in Kapitel acht, indem zusammenfassend eine Profilbildung der Charakteristika der drei untersuchten Bundesländer erstellt wird und dezidiert nochmals in Kapitel neun, in welchem die Ergebnisse unter drei leitenden Perspektiven interpretiert und diskutiert werden.

6.1 Ergebnisse in Bezug auf die quantitativen Aspekte

Wie viele Aufgabenstellungen werden durchschnittlich in einer schriftlichen Abiturprüfung gestellt und wie hoch ist der Anteil urteilsbildender Aufgaben?

Die Anzahl der Aufgabensets, welche aus dem Bundesland Thüringen in das Datenmaterial eingeflossen sind, betragen nur die Hälfte im Vergleich zu denen, die aus Niedersachsen und Bayern analysiert wurden. Dies hängt damit zusammen, dass in Bayern den Prüflingen stets vier verschiedene Wahlaufgabensets zur Verfügung stehen, während Thüringen und Niedersachsen nur zwei Wahlaufgaben anbieten. Die Anzahl der Aufgabensets ist in Niedersachsen dennoch ebenfalls doppelt so hoch wie in Thüringen, was auf die Unterscheidung zwischen erhöhtem und grundlegendem Anforderungsniveau zurückzuführen ist, welche in Thüringen und Bayern nicht gegeben ist. Daraus folgt, dass in den Bundesländern Bayern (sechs Jahre mit je vier Wahlaufgaben) und Niedersachsen (sechs

Jahre mit je zwei Wahlaufgaben mit je zwei Anforderungsniveaus) je 24 Aufgabensets im Untersuchungszeitraum eingesetzt wurden, in Thüringen hingegen nur zwölf (sechs Jahre mit je zwei Wahlaufgaben).

6.1.1 Die Gesamtaufgabenzahl

Im Sinne der oben genannten engen Aufgabendefinition (vgl. Kapitel 5.2.1), die eine Aufgabe als die kleinste zusammenhängende inhaltliche Handlungsaufforderung definiert, ergeben sich folgende *Gesamtaufgabenzahlen* im schriftlichen Abitur der Jahre 2014 bis 2019:

<i>Bundesland</i>	<i>Summe der Aufgabenstellungen 2014–2019</i>	<i>Aufgabensets 2014–2019</i>
<i>Bayern</i>	166	24
<i>Niedersachsen</i>	86 *	24
<i>Thüringen</i>	38	12
<i>Summe</i>	290	60

* (gA: 37, eA: 49)

Tab. 14: Summe der Aufgabenstellungen im schriftlichen Abitur des Faches Evangelische Religionslehre in den Jahren 2014–2019.

Somit wurden in dem Zeitraum von 2014 bis 2019 in den Bundesländern Bayern, Niedersachsen und Thüringen insgesamt 290 Aufgabenstellungen im schriftlichen Abitur des Faches Evangelische Religionslehre verteilt auf 60 Aufgabensets gestellt. Die obere Tabelle macht einen interessanten generellen Befund zu den Prüfungskonventionen der Bundesländer deutlich: Die Anzahl der Aufgabenstellungen im Verhältnis zu den Aufgabensets zeigt, dass die Prüflinge in Bayern im schriftlichen Abitur im Fach Evangelische Religionslehre mit rund 6,92 Aufgabenstellungen pro Aufgabenset mehr als doppelt so viele Aufgaben lösen mussten wie die Schülerinnen und Schüler in Thüringen, welche nur rund 3,17 Aufgaben für ein Prüfungsset zu bewältigen hatten. Niedersachsen liegt mit rund 4,08 Aufgaben pro Abitur im erhöhten Anforderungsniveau zwischen Bayern und Thüringen.¹

¹ Das erhöhte Anforderungsniveau für den Vergleich heran zu ziehen, ergibt daher Sinn, weil in Thüringen ausschließlich mit erhöhtem Anforderungsniveau geprüft wird und Bayern keine Unterscheidung zwischen erhöhtem und grundlegendem Anforderungsniveau vornimmt.

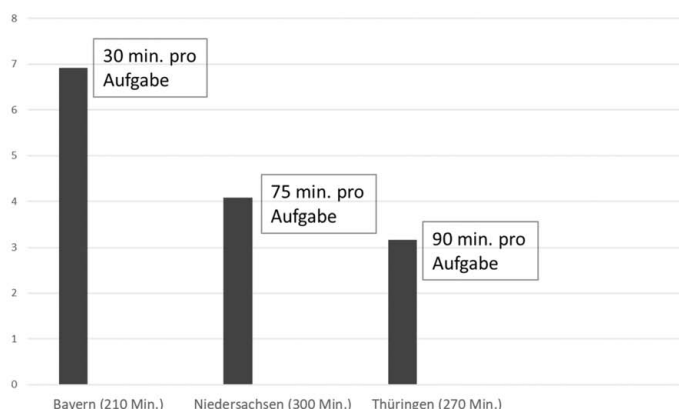


Abb. 11: Säulendiagramm zur durchschnittlichen Anzahl der Aufgabenstellungen je Aufgabenset (2014–2019) im schriftlichen Abitur im Fach Evangelische Religionslehre unter Angabe der Bearbeitungszeit.

6.1.2 Prozentualer Anteil ethischer Urteilsbildung

42 der 290 Aufgabenstellungen entsprechen der in Kapitel 5.2.2 diskutierten Definition:

Eine Aufgabenstellung wird dann in das Untersuchungsmaterial aufgenommen, wenn in ihr eine eigenständige Position im Sinne von Zustimmung oder Ablehnung hinsichtlich einer kontroversen ethischen oder religiösen Fragestellung vertreten werden soll.

Das Datenkorpus umfasst folglich 42 Fälle, welche sich folgendermaßen auf die drei Bundesländer verteilen:

Bundesland	Summe der Aufgabenstellungen zu ethischer und religiöser Urteilsbildung 2014–2019	Relative Häufigkeit der Urteilsbildung 2014–2019
Bayern	11 von 166	~ 7 %
Niedersachsen	20* von 86	~ 23 %
Thüringen	11 von 38	~ 29 %
gesamt	42 von 290	~ 14 %

* (gA: 11, eA: 9)

Tab. 15: Summe und relative Häufigkeit der Aufgabenstellungen zu ethischer und/oder religiöser Urteilskompetenz im schriftlichen Abitur des Faches Evangelische Religionslehre in den Jahren 2014–2019.

Betrachtet man die relative Häufigkeit von Aufgabenstellungen, die eine eigenständige ethische oder religiöse Urteilsbildung fordern, fallen diese in Bayern durch die hohe Gesamtaufgabenzahl, die ein Prüfling in einer schriftlichen Abiturprüfung im Fach Evangelische Religionslehre bewältigen muss, weitaus weniger ins Gewicht als das in Niedersachsen oder Thüringen der Fall ist.

Darüber hinaus sind neun Aufgabenstellungen im Untersuchungszeitraum gestellt worden, die eine Urteilsbildung aus einer vorgegebenen Perspektive fordern, welche jedoch nicht in das Untersuchungskorpus aufgenommen wurde, da es sich um keine eigenständigen Urteilsbildungen handelt (vgl. Kapitel 5.2.3). Von diesen neun Aufgaben stammen acht aus dem bayerischen Abitur.

Mindestens eine Aufgabenstellung zur eigenständigen Urteilsbildung ist in 60 % aller untersuchten Aufgabensets enthalten, wobei sich zwischen den Bundesländern prozentuale Unterschiede nachweisen lassen, wie die folgende Tabelle zeigt.

<i>Bundesland</i>	<i>Summe der Aufgabensets, die mind. eine ethische oder religiöser Urteilsbildung fordern 2014–2019</i>	<i>Relative Häufigkeit der Urteilsbildung 2014–2019</i>
<i>Bayern</i>	10 von 24	~ 42 %
<i>Niedersachsen</i>	18 von 24	75 %
<i>Thüringen</i>	8 von 12	~ 67 %
<i>Gesamt</i>	36 von 60	60 %

Tab. 16: Summe und relative Häufigkeit der Aufgabensets, die mindestens eine Aufgabenstellung zu ethischer und/oder religiöser Urteilskompetenz im schriftlichen Abitur des Faches Evangelische Religionslehre in den Jahren 2014–2019 aufweisen.

6.2 Ergebnisse in Bezug auf die Aufgabenstellungen

Wie lassen sich die Aufgabenstellungen, welche Urteilsfähigkeit prüfen, im untersuchten Prüfungszeitraum beschreiben?

6.2.1 Aufgabenart (A)

Urteilsbildung wird in allen drei Bundesländern im Untersuchungszeitraum in der Regel mit einem Textbezug geprüft. Ausnahmen bilden nur zwei Aufgabenstellungen aus Bayern, bei denen die Prüflinge über eine Fragestellung urteilen sollten, ohne dass dem Aufgabenset ein die Fragestellung aufgreifender

Text zugrunde lag [BY, 2015 (3) und BY, 2019 (4)]². Unter den 42 analysierten Aufgabenstellungen zur Urteilsfähigkeit, finden sich nur zwei, die als »erweiterte Textaufgabe«³ zu kategorisieren sind, die also auch nichttextliches Material einbeziehen. Es handelt sich zum einen um ein Titelbild des Magazins DER SPIEGEL, welches in Niedersachsen 2014 Grundlage einer Fragestellung war [NI, 2014, eA (1)], und zum anderen um eine Abbildung des Freskos *Invidia* (Neid) von Giotto, welches in einer Aufgabenstellung in Bayern Teil des einzubeziehenden Materials war [BY, 2017 (3)].⁴ Überwiegend – in 22 Fällen – sind Urteilsbildungsaufgaben als Textaufgaben zu klassifizieren, also als Aufgaben, in denen eine Auseinandersetzung mit einem gegebenen Text erfolgen soll, ohne dass diese Auseinandersetzung nochmals kontextualisiert oder adressatengebunden ist. Einen mit 17 von 42 Fällen ebenso beachtlichen Anteil stellen die – in der Aufgabenforschung als innovativ geltenden – sogenannten Gestaltungsaufgaben dar, die sich durch eine adressatengebundene Produktorientierung auszeichnen.⁵ Während in Bayern und Thüringen die Gestaltungsaufgaben über die Hälfte der Aufgaben zur Urteilsbildung ausmachen (~55%; jeweils 6 von 11 Aufgaben), kommen diese in Niedersachsen nur in einem Viertel der urteilsbildenden Aufgabenstellungen zum Einsatz (5 von 20).

6.2.2 Fragestellungen der Urteilsbildung (B)

Die Fragestellungen der Urteilsbildung im schriftlichen Abitur sind vielfältig. Themen, die mehrfach im Kontext mit Urteilsbildungen auftauchen, sind Gottesbilder, Religionskritik, Grenzen und Chancen des technisch Machbaren, Leiderfahrungen und Sterbehilfe oder der Einsatz von Gewalt. 23 der insgesamt 42 Fragestellungen sind dabei als ethisch-religiöse Fragestellungen zu kategorisieren bzw. besitzen eine unmittelbare Verknüpfung mit dem Thema Religion (beispielsweise ein Urteil über die Bedeutung von Gottesbeweisen für den Gottesglauben [TH, 2018, (2)]). Dies entspricht ~55 %. 19 Fragestellungen weisen keinen dezidiert religiösen Bezug auf (~45 %). Es handelt sich um ethisch-phi-

² In beiden Fällen liegt dem Aufgabenset ein Text zugrunde, der jedoch nicht die zu beurteilende Fragestellung aufgreift. Das Aufgabenset [TH, 2014 (2)] besteht aus zwei Handlungsaufforderungen, wobei nur die erste einen expliziten Bezug zum Material herstellt. Da das Material jedoch ebenso auf die zweite Aufgabenstellung passt und die Handlungsaufforderungen nicht durch Nummerierungen o.ä. klar voneinander getrennt sind, ist diese Aufgabe nicht losgelöst von den Materialien zu betrachten.

³ Die Kategorisierung leitet sich aus den EPA-Vorgaben ab (vgl. Kap. 5.3.2 Kategorie A).

⁴ Die Analyse und Interpretation des Freskos wurden in einer anderen Aufgabenstellung gefordert. In die Urteilsbildung sollten die Ergebnisse der Interpretation einbezogen werden.

⁵ Vgl. OTTO/KÜHN, Zwischen Tradition und Innovation, 6.

Aufgabenart Urteilsbildung fordernder Aufgabenstellungen

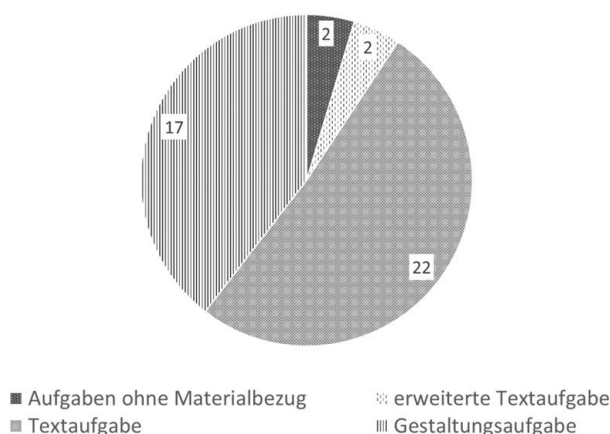


Abb. 12: Zusammenfassendes Kreisdiagramm zu den Aufgabenarten urteilsbildender Aufgabenstellungen von 2014–2019 in den Bundesländern Bayern, Niedersachsen und Thüringen (n = 42).

losophische Fragestellungen, wie beispielsweise nach der Umsetzbarkeit des Pressecodex [BY, 2017 (1)], die ihren Bezug zur Religion über Reflexionsanforderungen erhalten (im genannten Beispiel: »[...] Beziehen Sie dabei [bei der Erörterung der Fragestellung; Anm. K.M.] grundlegende Aspekte philosophischer und christlicher Ethik mit ein! [BY, 2017 (1)]«). Mit Blick auf die Fälle der Studie halten sich also ethisch-religiöse Fragestellungen in etwa die Waage mit Fragestellungen ohne dezidiert religiöse Bezüge.⁶ Das Verhältnis zwischen durch konkrete Positionen vermittelten Fragestellungen und Fragestellungen ohne eine vermittelnde Position ist ebenso recht ausgewogen (18:24), wie folgende Kreuztabelle zeigt:

⁶ Zur Kategorisierung der Fragestellungen in ethisch-religiös und ethisch-philosophisch siehe Kapitel 2.1.

Ausrichtung der Fragestellung			Summe
Art der Vermittlung (mit Position ver- knüpft oder nicht)	ethisch-religiöse Fragestellung (9)	ethisch-philosophisch Fragestellung (9)	18 ohne konkrete Position
	theologische Position zu ethisch-religiöser Fragestellung (13)	theologische Position zu ethisch-philosophischer Fragestellung (10)	24 mit konkreter Position
	philosophische Positi- on zu ethisch-religiöser Fragestellung (1)		
Summe	23 ethisch-religiöse Fragestellung	19 ethisch-philosophische Fragestellung	Gesamt 42

Tab. 17: Systematisierung der Fragestellungen der Urteilsbildung.

Auch wenn sich die Verteilung auf die vier Bereiche (ethisch-religiös/ethisch-philosophisch und vermittelt/nicht vermittelt) mit Blick auf die Gesamtzahl der Aufgabenstellung als sehr ausgewogen darstellt, sind mit Blick auf die einzelnen Bundesländer eindeutige Schwerpunktsetzungen zu erkennen, welche im Folgenden gezeigt werden sollen.

Bayern

In Bayern wurden im Untersuchungszeitraum acht Urteilsbildungen zu ethisch-philosophischen Fragestellungen gefordert, die nicht genuin auch religiöse Fragestellungen sind. Zwei Fragestellungen gewinnen einen religiösen Bezug dadurch, dass über eine theologische Position geurteilt werden soll, wobei eines der Themen einen dezidiert religiösen Bezug aufweist (Verhältnis von Kirche und Staat), während der religiöse Bezug des anderen Themas zum Einsatz von Gewalt durch die Perspektive der vorgegebenen Position ins Spiel kommt (eine Positionierung eines Theologen zum Einsatz von Gewalt). Eine Fragestellung nach der Bedeutung des Apostolikums für die Beantwortung »letzter Fragen« lässt sich als dezidiert religiöse Fragestellung beschreiben.

Die Fragestellungen zur Urteilsbildung in Bayern sind damit in erster Linie ethisch-philosophischer Art. Die religiösen Bezüge kommen nicht durch die Themenwahl, sondern durch Reflexionsaufforderungen zustande, wie im folgenden Kapitel gezeigt werden wird. Die Tabelle gibt Auskunft über die Fragestellungen der Urteilsbildung:

Fragestellungen der Urteilsbildung in Bayern (n = 11)
Ethische Fragestellung ohne dezidiert religiösen Bezug (8)
- Umgang mit Doping und Selbstoptimierung im Sport [BY, 2014 (2)] - Sinn von Vorbildern für die Lebensgestaltung [BY, 2014 (4)] - Demut als schützenswerter Wert [BY, 2015 (3)] - Dankbarkeit und Demut trotz Leiderfahrungen⁷ [BY, 2015 (3)] - Umsetzbarkeit des Pressecodex [BY, 2017 (1)] - Umgang mit Neid [BY, 2017 (3)] - Kryonik⁸ [BY, 2017 (4)] - Chancen und Grenzen des technisch Machbaren an selbstgewähltem Beispiel [BY, 2018 (4)]
Religiöse Fragestellung (1)
- Antworten des Apostolikums auf »letzte Fragen« [BY, 2019 (4)]
Religiöse Perspektive auf Fragestellung ohne dezidiert religiösen Bezug (1)
- konkrete Perspektive eines ev. Theologen zum Einsatz von Gewalt [BY, 2016 (4)]
Religiöse Perspektive auf religiöse Fragestellung (1)
- frei gewähltes Thema aus einem Interview mit einem ev. Theologen zum Verhältnis von Kirche und Staat [BY, 2015 (4)]
Nicht-religiöse Perspektive auf eine religiöse Fragestellung (0)
- keine

Tab. 18: Fragestellungen der Urteilsbildung in Bayern im Zeitraum von 2014–2019.

Niedersachsen

Anders als in Bayern sind in Niedersachsen mit Ausnahme einer Aufgabenstellung konkrete Positionierungen Gegenstände der Urteilsbildung. In 18 von 19 Fällen ist die vorgegebene Position als christlich zu kategorisieren. In einem der 19 Fälle handelt es sich um eine nicht-religiöse Position, die beurteilt werden soll. Die Themen, zu denen eine christliche Position vorgegeben wird, sind in zehn von 18 Fällen religiöse Themen und in acht Fällen ohne dezidiert religiösen Bezug (beispielsweise Präimplantationsdiagnostik, Umgang mit Geflüchteten, Menschenbild und Freiheitsverständnis).

⁷ Die Frage nach der Dankbarkeit wird im Material (eine Leseranfrage aus dem Magazin der Süddeutschen Zeitung) ohne dezidierten Bezug auf eine Dankbarkeit gegenüber Gott diskutiert.

⁸ Kryonik bezeichnet ein Verfahren, bei dem Organismen oder einzelne Organe nach dem Tod durch Einfrieren in Stickstoff konserviert werden, um sie zu einem späteren Zeitpunkt – wenn der medizinische Fortschritt es möglich macht – reanimieren zu können.

Der Ausnahmefall, in dem keine konkrete Positionierung Urteilsgegenstand ist, fordert eine Auseinandersetzung mit einem Titelbild des SPIEGEL-Magazins, welches ein Kräftering des islamischen und des christlichen Gottes anhand zweier Hände darstellt. Auch dieser Aufgabenstellung liegt ein theologischer Sachtext zugrunde, der zwar berücksichtigt werden soll, jedoch nicht Urteilsgegenstand ist.

Hervorzuheben ist darüber hinaus, dass unter den zehn konkret christlichen Perspektiven auf religiöse Fragestellungen sechs katholische Perspektiven beurteilt werden sollen. Mit dem Positionspapier der Arbeitsgemeinschaft Christlicher Kirchen in Bayern kommt zudem eine ökumenische Perspektive auf eine ethisch-philosophische Fragestellung ins Spiel. Die Positionen treten jedoch nicht als spezifisch katholisch in den Blick. Der Hinweis zur Konfessionalität der zu diskutierenden Positionen wird den Prüflingen zwar gegeben, er wird jedoch – auch in den Erwartungshorizonten – nicht weiter aufgegriffen. Die Kontroversen um religiöse Fragestellungen werden folglich in einen überkonfessionellen Kontext gestellt.⁹

Fragestellungen der Urteilsbildung in Niedersachsen (n = 20)

Ethische Fragestellung ohne dezidiert religiösen Bezug

- keine

Religiöse Fragestellung (1)

- Titelblatt des »SPIEGEL« zum Verhältnis der Gottesbilder in Christentum und Islam [NI, 2014, eA (1)]

Religiöse Perspektive auf Fragestellung ohne dezidiert religiösen Bezug (8)

- konkrete Position eines ev. Theologen zur Präimplantationsdiagnostik [NI, 2014, gA (1)]

⁹ Der in Kapitel 4.2.1 zitierte Befund BERND SCHRÖDERS, dass in niedersächsischen Kerncurricula keine konfessionellen Differenzierungen hinsichtlich ethischer Fragestellungen zu finden sind, lässt sich folglich auch an den Abituraufgabenstellungen zeigen (vgl. SCHRÖDER, Was macht es für einen Unterschied, 41–63). Die von SCHRÖDER geforderte konfessionelle Ausdifferenzierung findet nicht statt. Dies ist vor allem auch vor dem Hintergrund aktueller Diskussionen über Konfessionssensibilität im Kontext des konfessionell-koooperativen Religionsunterrichts interessant: Mit einer Betonung der Differenzen würde man die konfessionellen Profile stärken und greifbar machen, während man mit dem Fokus auf das Gemeinsame die Ökumene stärkt. Im vorliegenden Beispiel spielen die konfessionellen Hintergründe der Positionen jedoch schlichtweg keine Rolle.

¹⁰ Diese Aufgabenstellung wurde 2019 sowohl in der Prüfung mit erhöhtem als auch in der Prüfung mit grundlegendem Anforderungsniveau gestellt, so dass sie hier doppelt aufgeführt wird.

-
- konkrete Position der Arbeitsgemeinschaft Christlicher Kirchen in Bayern zum Umgang mit Geflüchteten [NI, 2015, eA (2)]
 - konkrete Position eines ev. Theologen zum Menschenbild [NI, 2016, gA (1)]
 - konkrete Position eines ev. Theologen zum Freiheitsverständnis [NI, 2016, eA (1)]
 - konkrete Position eines Christen zur Sterbehilfe [NI, 2018, gA (1)]
 - konkrete Position eines ev. Theologen zu Sterben und Selbstbestimmung [NI 2018 eA (1)]
 - konkrete Position einer ev. Theologin zum Einsatz von Gewalt gegen Terror [NI, 2019 gA (1)]
 - konkrete Position einer ev. Theologin zum Einsatz von Gewalt gegen Terror¹⁰ [NI, 2019 eA (1)]
-

Religiöse Perspektive auf religiöse Fragestellung (10)

- konkreter Ansatz einer kath. Theologin zum Dialog zwischen Christen und Muslimen [NI, 2014, gA (2)]
 - konkrete Position zweier ev. Theologen zur Theodizee-Frage [NI, 2015, gA (2)]
 - konkrete Position eines kath. Theologen zur Theodizee-Frage [NI, 2015, eA (1)]
 - Zitat eines kath. Theologen über die Bedeutung der Religionskritik [NI, 2016, gA (2)]
 - konkrete Position eines ev. Theologen zu Sünde und Vergebung/Gott und Mensch [NI, 2017, gA (1)]
 - konkrete Position eines ev. Theologen zu Sünde und Vergebung/Gott und Mensch¹¹ [NI, 2017, eA (2)]
 - das Konzept einer »prophetischen Kirche« eines altkath. Theologen [NI, 2017, gA (2)]
 - konkrete Position eines kath. Theologen zu religionskritischen Einwänden [NI 2018 gA (2)]
 - konkrete Position eines ev. Theologen zu Gottesbildern [NI 2019 gA (2)]
 - konkrete Position eines kath. Theologen zur Frage, ob Gott in Versuchung führt [NI 2019 eA (2)]
-

Nicht-religiöse Perspektive auf eine religiöse Fragestellung (1)

- Nicht-religiöse Perspektive auf die Frage nach der Rolle der Religionen in der Zukunft [NI 2018 eA (2)]
-

Tab. 19: Fragestellungen der Urteilsbildung in Niedersachsen im Zeitraum von 2014–2019.

Thüringen

In Thüringen sind die Fragestellungen ebenso überwiegend als *religiöse Fragestellungen* zu definieren, jedoch sind diese – bis auf zwei Ausnahmen – nicht an konkrete christliche Positionen geknüpft, sondern durch die *Inhalte der Fragestellung* religiös. Neun der insgesamt elf Fragestellungen fordern eine Urteils-

¹¹ Auch diese Aufgabenstellung wurde sowohl in der Prüfung mit erhöhtem als auch in der Prüfung mit grundlegendem Anforderungsniveau gestellt, so dass auch sie hier doppelt aufgeführt wird.

bildung über ein religiöses Thema bzw. eine theologische Position, während nur eine Fragestellung im Untersuchungszeitraum als ethische Fragestellung ohne dezidierten theologischen Bezug definiert werden kann (verantwortungsvolle Gestaltung von Freiheit). Eine weitere Fragestellung hat eine christliche Position zu einer ethischen Fragestellung ohne dezidiert religiösen Bezug (Präimplantationsdiagnostik) zum Urteilsgegenstand.

Fragestellungen der Urteilsbildung in Thüringen (n = 11)
Ethische Fragestellung ohne dezidiert religiösen Bezug (1)
– verantwortungsvolle Gestaltung von Freiheit [TH, 2017 (1)]
Religiöse Fragestellung (7)
– Botschaft Jesu [TH, 2014 (2)]
– Rolle christlicher Maßstäbe für persönliche Entscheidungen [TH, 2014 (2)]
– Sterben und Tod in Auseinandersetzung mit der Theodizee-Frage [TH, 2016 (1)]
– heutige Relevanz des lutherischen Freiheitsverständnisses [TH, 2017 (1)]
– Relevanz der Bergpredigt für die Politik [TH, 2017 (2)]
– Individualität persönlicher Glaubensbekenntnisse [TH, 2018, (1)]
– Bedeutung von Gottesbeweisen für den Gottesglauben [TH, 2018, (2)]
Religiöse Perspektive auf Fragestellung ohne dezidiert religiösen Bezug (1)
– Position der ev. Kirche zur Präimplantationsdiagnostik [TH, 2015 (2)]
Religiöse Perspektive auf religiöse Fragestellung (2)
– Konkrete Position einer ev. Theologin zur politischen Verantwortung von Christen [TH, 2014 (1)]
– Gottesbilder in einem konkreten Glaubensbekenntnis [TH, 2019, (1)]
Nicht-religiöse Perspektive auf eine religiöse Fragestellung (0)
– keine

Tab. 20: Fragestellungen der Urteilsbildung in Thüringen im Zeitraum von 2014–2019.

Vergleichende Perspektiven

In den folgenden Kreisdiagrammen werden die Fragestellungen der Urteilsbildung in den einzelnen Bundesländern nochmals miteinander verglichen, wobei folgende Kategorien unterschieden werden:

- 1.) ethische Fragestellung ohne dezidiert religiösen Bezug
- 2.) religiöse Fragestellung
- 3.) theologische Position zu einer ethischen Fragestellung ohne dezidiert religiösen Bezug sowie
- 4.) theologische Position zu einer religiösen Fragestellung.

Dabei ist zu betonen, dass an dieser Stelle ausschließlich die Fragestellungen der Urteilsbildung in den Blick genommen werden, so dass in der Regel auch Aufgabenstellungen, deren Urteilsgegenstand keinen dezidiert religiösen Bezug aufweist, durch Reflexionsaufforderungen, wie sie im folgenden Kapitel beschrieben sind, religiöse Perspektiven mitbedenken.



Abb. 13: Kreisdiagramm über die Fragestellungen der Urteilsbildung in Bayern.



Abb. 14: Kreisdiagramm über die Fragestellungen der Urteilsbildung in Niedersachsen.

Gegenstände der Urteilsbildung in Thüringen



Abb. 15: Kreisdiagramm über die Fragestellungen der Urteilsbildung in Thüringen.

Ein Vergleich macht deutlich, dass es unterschiedliche Schwerpunktsetzungen und Stile zwischen den verglichenen Bundesländern im Untersuchungszeitraum gibt: Während im bayerischen Abitur in erster Linie Urteilsfähigkeit an ethischen Fragestellungen ohne dezidiert religiöse Bezüge unvermittelt geprüft wird, stehen im thüringischen Abitur mit sieben von elf Aufgaben dieser Kategorie vor allem religiöse Fragestellungen im Zentrum von Urteilsbildungsaufgaben. Niedersachsen hingegen legt einen hohen Wert darauf, dass eine Position vermittelt auftritt: 18 der 20 aus Niedersachsen analysierten Aufgabenstellungen machen die Position einer Theologin oder eines Theologen zum Urteilsgegenstand. In Bayern tritt diese Form der Fragestellung nur in zwei Fällen und in Thüringen in drei Fällen von jeweils elf analysierten Aufgabenstellungen auf.

6.2.3 Reflexionskriterien (C)

Im Hinblick auf die Reflexionskriterien sind bereits die induktiv entstandenen Subkategorien des Kategoriensystems ein interessantes Ergebnis der Analyse, da an ihnen deutlich wird, welcher Art die Reflexionsanforderungen im Untersuchungszeitraum waren: Christliche sowie nicht-christliche Sichtweisen sollten reflektiert werden, wobei die gegebenen Materialien aber auch eigene Erfahrungen einzubeziehen waren. Hinzu kommen Kontextualisierungen der Aufgabenstellungen, die einen spezifischen Rahmen vorgeben und dadurch ebenso einen Reflexionsanlass bieten, da die eigene Urteilsbildung reflektiert werden muss und Argumente gegebenenfalls an den Rahmen angepasst werden müssen.

In 25 der 42 Aufgaben des Untersuchungskorpus werden keine spezifischen Reflexionsanforderungen gestellt. In den 17 Aufgabenstellungen, die spezifische Reflexionen im Urteilsbildungsprozess fordern, kommen mitunter mehrfache Reflexionsanforderungen vor. In zehn der 17 Aufgabenstellungen wurde der Einbezug der Materialien gefordert. Über die Art des Materials gibt Kategorie D Auskunft (vgl. Kapitel 6.2.4).

Ebenfalls in zehn der 17 Aufgaben mit Reflexionsanforderungen sollte die Urteilsbildung an einer christlichen Weltsicht reflektiert werden wie beispielsweise durch »Grundzüge biblischer Ethik« [BY, 2016 (4)], »Aspekte biblischer Anthropologie und Ethik« [BY, 2017 (3)] oder durch »evangelische Einsichten zu Gott und Mensch« [BY, 2014 (4)]. Dabei wird am häufigsten eine Reflexion der Urteilsbildung an einem christlichen/biblischen/evangelischen Menschenbild gefordert (fünfmal). Die Reflexion an einer biblischen/theologischen/evangelischen (Sozial-)Ethik wird viermal genannt. Zweimal wird die Reflexion anhand geeigneter Texte aus der Bibel gefordert. Jeweils einmalige Nennungen sind ein evangelisches Gottesbild, christliche Zukunftshoffnung, theologische Angemessenheit und Denkmodelle zum politischen Engagement von Christen.

Die anderen Reflexionsanforderungen kommen selten vor: In zwei Fällen ist der Einbezug einer nicht-christlichen Sicht gefordert (der Einbezug eines religionskritischen Konzepts [NI, 2016, gA (2)] sowie Aspekte philosophischer Ethik [BY, 2017 (1)]). In ebenso zwei Fällen entsteht eine Reflexion durch den Kontext (es soll ein Forenbeitrag für »evangelisch.de« verfasst [NI, 2014, gA (1)] sowie ein Beitrag für die Schülerzeitung für das Fach Evangelische Religionslehre geschrieben werden [BY, 2018 (4)]). In einem Fall wird der Einbezug persönlicher Erfahrungen gefordert [TH, 2014 (1)].

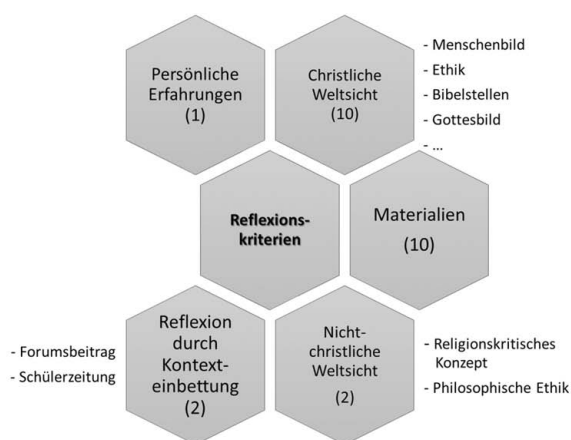


Abb. 16: Reflexionskriterien der Urteilsbildung im Untersuchungszeitraum 2014–2019.

6.2.4 Exkurs: Rolle christlicher Sichtweisen in den Aufgabenstellungen zur Urteilsfähigkeit

Besonders interessant ist die Analyse der christlichen Reflexionsanforderungen in Bezug auf die neun Aufgabenstellungen, deren Fragestellungen keinen dezidiert religiösen Bezug besitzen und die auch nicht durch eine theologische Sicht auf die Fragestellung einen religiösen Bezug erhalten (vgl. Kategorie B Merkmalsausprägung B1): In den meisten dieser Fragestellungen – die überwiegend aus Bayern kommen – wird der Bezug zur Religion über eine christliche Reflexionsanforderung hergestellt (in sieben der neun Fälle). Auch die zwei Aufgabenstellungen, in denen dies nicht der Fall ist, stellen einen Bezug zu einer christlichen Sicht her: Eine Aufgabenstellung durch die Kontextualisierung der Aufgabenstellung (der Schülerzeitungsartikel für das Fach Evangelische Religionslehre [BY, 2018 (4)]), eine andere durch die Forderung, die Materialien einzubeziehen, wobei ein Material eine christliche Sichtweise ins Spiel bringt. Somit stellen alle im Untersuchungszeitraum gestellten Urteilsbildungsaufgabenstellungen einen Bezug zur Religion her, auch wenn nicht alle Fragestellungen dezidiert religiös sind.

Bayern

Wie die Analyse der Fragestellungen urteilsbildender Aufgaben in der vorangegangenen Kategorie gezeigt hat, ist Bayern das Bundesland mit den meisten Fragestellungen ohne dezidiert religiösen Bezug. Dafür stellt Bayern im Gegensatz zu den anderen beiden untersuchten Bundesländern am häufigsten spezifische Reflexionsanforderungen. Nur in zwei der 11 Aufgabenstellungen aus Bayern sind keine speziellen Reflexionsanforderungen zu finden ([BY, 2015 (4)], [BY, 2019 (4)]). Siebenmal wird der Einbezug christlicher Sichtweisen gefordert, fünfmal der Einbezug von Materialien, einmal der Einbezug philosophischer Ethik und in einem Fall stellt die Kontexteinbettung der Aufgabenstellung die Reflexionsanforderung her.

Niedersachsen

In Niedersachsen tauchen im Gegensatz zu Bayern nur selten explizite Reflexionsanforderungen in den Aufgabenformulierungen auf: Nur vier der 20 Aufgabenstellungen fordern eine spezifische Reflexion. In zwei dieser Fälle soll das Material in den Urteilsbildungsprozess einbezogen werden. Eine Aufgabenstellung fordert eine religionskritische Reflexion und eine weitere Aufgabenstellung gibt durch die Kontexteinbettung einen Reflexionsimpuls. In keiner der Aufgabenstellungen wird eine christliche Reflexionsanforderung formuliert. Dies kann durch die Art der Urteilsbildungsaufgaben in Niedersachsen erklärt werden: Wie in Kategorie B zu den Fragestellungen der Urteilsbildung dargestellt, kommt in Niedersachsen keine Fragestellung ohne dezidiert religiösen Bezug vor. In den

meisten Fällen ist entweder die Fragestellung religiös oder eine religiöse Position auf eine ethische Fragestellung soll beurteilt werden. Christliche Sichtweisen sind in Niedersachsen folglich nicht Reflexionskategorie, sondern Urteilsgegenstand.

Thüringen

In Thüringen ist das Verhältnis zwischen Urteilsbildungen ohne spezifische Reflexionsanforderungen und denen mit spezifischen Reflexionsanforderungen recht ausgewogen: In sechs Aufgabenstellungen sind keine spezifischen Reflexionsanforderungen gestellt. In den fünf Aufgabenstellungen mit Reflexionsanforderung kommt dreimal die Forderung vor, christliche Perspektiven einzubeziehen, dreimal soll das Material einbezogen werden und in einem Fall sollen persönliche Erfahrungen bedacht werden.

6.2.5 Textsorten der Materialien, die den Aufgaben zur Urteilsbildung zugrunde liegen (D)

Urteilsbildungsaufgaben im schriftlichen Abitur im Fach Evangelische Religionslehre werden in der Regel in Zusammenhang mit Materialien gestellt: Wie bereits in der Kategorie *Aufgabenart* erwähnt, spielen in nur zwei Aufgabenstellungen des Datenkorpus' Materialien im Urteilsbildungsprozess keine Rolle (vgl. Kapitel 6.2.1). Eine Statistik und ein Zitat sind jeweils nur einmal im Untersuchungszeitraum als Material für Urteilsbildungsaufgabenstellungen zu finden. Ein Bild soll in zwei Fällen einbezogen werden. Auffällig ist darüber hinaus, dass Bibelverse und -perikopen ausschließlich in Thüringen als Materialien vorkommen. Insgesamt lässt der Materialeinsatz deutliche Präferenzen der Bundesländer erkennen: Während Thüringen vielfach auch lyrische Texte sowie die genannten Bibelverse und -perikopen als Materialgrundlage für Urteilsbildungsaufgaben zur Verfügung stellt, tauchen diese beiden Textsorten in Bayern und Niedersachsen gar nicht auf. Niedersachsen legt einen deutlichen Schwerpunkt auf Sachtexte, während Bayern sowohl Sachtexte als auch informierende Zeitungsartikel als Materialgrundlage zur Verfügung stellt. Im Folgenden wird die Analyse der einzelnen Bundesländer noch vertieft.

Ein weiterer interessanter Befund im Vergleich der Bundesländer ist der, dass in den Abiturprüfungen Niedersachsens – bis auf drei Ausnahmen – ein Material pro Aufgabenset zur Verfügung gestellt werden (im Durchschnitt ergibt dies 1,15 Materialien pro Aufgabenstellung). Thüringen legt seinen Aufgabenstellungen bis zu fünf verschiedene Materialien zugrunde (im Durchschnitt 2,45), während Bayern in der Regel ein Material, in drei Fällen jedoch auch vier Materialien zur Verfügung stellt (Durchschnitt 1,64). Dies sagt nichts über die Länge der zu analysierenden Materialien aus, da die Sachtexte in Niedersachsen meist

mehrere Seiten umfassen, während beispielsweise in Thüringen auch sehr kurze Textauszüge oder Gedichte ein Material darstellen.

Bayern

In den Materialien aus bayerischen Abiturprüfungen sind im Vergleich zu den anderen beiden Bundesländern die meisten informierenden Texte (Nachrichten, Berichte, Reportagen o.ä.) zu finden (in sechs Fällen). Diese halten sich mit den ethischen und/oder theologischen Sachtexten, die in sieben Aufgabenstellungen eine Rolle spielen, nahezu die Waage. Meinungsäußernde Texte, wie Forenbeiträge im Internet oder Leserkommentare, kommen in drei Fällen vor. In je einem Fall ist ein Bild bzw. eine Statistik aufgegriffen, woraus sich ergibt, dass die Textsorten der zugrunde liegenden Materialien sehr breit gefächert sind. Wenngleich den Prüflingen zur Lösung der Aufgabenstellungen, wie in den anderen untersuchten Bundesländern auch, eine Bibel zur Verfügung steht, ist die Bibel nur indirekt – in Form einer Berücksichtigung des christlichen Menschenbildes beispielsweise – eine Bezugsgröße für die Urteilsbildung. Den Prüflingen stehen Auswahl und Nutzung von Bibelstellen somit frei.

Niedersachsen

Niedersachsen setzt einen deutlichen Schwerpunkt auf Sachtexte, welche in der Regel theologischer Art sind. Den 20 Aufgabenstellungen liegen 21 ethische und/oder theologische Sachtexte zugrunde. In nur einem Fall kommt ein in erster Linie meinungsäußernder Text in Form eines Leserbriefs zum Einsatz, der jedoch auch an einen Sachtext gekoppelt ist und mit diesem zusammen die Materialgrundlage der Aufgabenstellung bildet [NI, 2018, gA (1)]. In wiederum einem Fall ist ein Bild Grundlage einer Urteilsbildungsaufgabe. Die Bibel spielt auch hier in den Materialien keine Rolle, auch wenn sie als Hilfsmittel zur Lösung der Aufgaben zur Verfügung steht.

Thüringen

Die Materialien im Bundesland Thüringen fallen durch die Häufung lyrischer Texte auf (Gedichte, Liedtexte, Gebete, persönliche Bekenntnisse), derer zwölf den insgesamt elf Aufgabenstellungen zugrunde liegen. Damit bildet diese Kategorie in Thüringen die stärkste Gruppe. Bibelverse und -perikopen sind sechsmal Materialgrundlage – was ebenso ein Alleinstellungsmerkmal Thüringens darstellt.¹² Als Sachtexte zu klassifizierende Texte kommen fünfmal vor.

¹² Übereinstimmend damit wird in den Erwartungshorizonten wiederholt als Qualitätsmerkmal genannt, dass die »Aussagen mit geeigneten Textstellen aus der Bibel« gestützt sein sollen. Dabei sind nur in einigen Fällen Bibelstellen als Materialien zu den Aufgaben angegeben. In zwei Fällen sind keine Bibelstellen vorgegeben, ihr Einbezug wird jedoch verbindlich verlangt. In zwei weiteren Fällen wird der Einbezug »gegebenenfalls« verlangt. Von

Eher selten sind informierende Texte wie Nachrichten oder Berichte, derer drei den Materialien zugrunde liegen. In einem Fall kommt ein Zitat zum Einsatz. Thüringen setzt wiederholt mehrere kurze Materialien pro Aufgabenset ein, was sich von dem Vorgehen in Niedersachsen – mit jeweils nur einem längeren Text – unterscheidet.

6.3 Ergebnisse in Bezug auf die Erwartungshorizonte

Welche Bewertungshinweise erhalten korrigierende Lehrkräfte durch die Erwartungshorizonte hinsichtlich der Einschätzung von Urteilsbildungsaufgaben?

6.3.1 Passung zwischen geforderter und erwarteter Urteilsfähigkeit (E)

Bevor die obige, für dieses Kapitel leitende Fragestellung beantwortet wird, soll ein Fokus auf die Passung zwischen den Forderungen nach Urteilsbildungen in den Aufgabenstellungen und den Erwartungshorizonten gelegt werden. Dazu wurden nicht nur die Aufgabenstellungen, welche nach der in Kapitel 5.2.2 dargestellten Definition Urteilsfähigkeit prüfen, analysiert, sondern die Erwartungshorizonte aller im Untersuchungszeitraum gestellten Aufgabenstellungen hinsichtlich ihrer Forderung einer eigenständigen Positionierung. Auf diese Weise ist sichergestellt, dass auch jene Unstimmigkeiten erfasst werden, in denen die Aufgabenstellung keine Urteilsbildung fordert, die Erwartungshorizonte diese aber verlangen. In den Erwartungshorizonten zu den 290 Aufgabenstellungen wurde in 41 Aufgabenstellungen eine ethische oder religiöse Urteilsbildung gefordert. Dies weicht zwar nur um einen Fall von der Summe der in die Studie aufgenommenen Fälle ab, jedoch sind die Aufgabenstellungen nicht deckungsgleich: Neben Niedersachsen, in denen die Anforderungen in den Aufgabenstellungen mit den Forderungen der Erwartungshorizonte in allen 20 Fällen übereinstimmen, gibt es in den Aufgabenstellungen aus Bayern und Thüringen zum einen Aufgabenstellungen, die nach der Definition in Kapitel 5.2.2 eine Urteilsbildung fordern, diese in den Erwartungshorizonten aber nicht einfordern und konkretisieren. Zum anderen gibt es Aufgabenstellungen, die in ihrer Formulierung keine eigenständige Urteilsbildung verlangen, während der Erwartungshorizont eine Urteilsbildung fordert. Diese im Folgenden genauer beschriebenen Unstimmigkeiten kamen in Bayern in elf und in Thüringen in drei Aufgabenstellungen vor.

den Prüflingen wird folglich eine grundlegende Kenntnis zentraler Bibelstellen erwartet, wenngleich ihnen – wie in den anderen Bundesländern auch – die Bibel als Hilfsmittel zur Verfügung steht.

Bayern

Die Durchsicht aller Erwartungshorizonte der 166 in Bayern im Untersuchungszeitraum gestellten Aufgabenstellungen zeigt, dass eine eigenständige Urteilsbildung in neun Fällen in den Erwartungshorizonten angesprochen wird (Stichworte: »Stellungnahme« [BY, 2015 (3)], »eigene Antwort« auf die gegebene Fragestellung [BY, 2015 (3)], »Abwägung und Begründung durch die Schülerinnen und Schüler« [BY, 2015 (2)], »Stellungnahme« [BY, 2015 (4)], »begründetes eigenes Urteil« [BY, 2016 (4) und BY, 2019 (4)], eine Darstellung »bewerten [und diskutieren]« [BY, 2016 (3)], ein »begründetes und differenziertes abschließendes Urteil« [BY, 2019 (2)] und »eine differenzierte Positionierung« [BY, 2019 (4)]).¹³

In Bezug auf die aus Bayern stammenden Aufgabenstellungen liegt eine hohe Unstimmigkeit zwischen den Forderungen im Erwartungshorizont und der zugrunde liegenden Definition von Urteilsfähigkeit vor. Von den neun im Erwartungshorizont geforderten Urteilsbildungen sind entsprechend der Definition ethischer und religiöser Urteilsfähigkeit in Kapitel 5.2.2 nur fünf Aufgabenstellungen in das Untersuchungskorpus aufgenommen worden. Gründe für den Ausschluss waren in einem Fall ein von der Aufgabe geforderter Perspektivwechsel: Die Prüflinge sollten nicht aus eigener, sondern aus einer anderen, hier jeweils christlichen Perspektive urteilen, wobei die Perspektivübernahme in einem der Erwartungshorizonte keine Erwähnung mehr findet und stattdessen die Urteilsbildung des Prüflings als Kriterium angelegt wird, so dass eine Unstimmigkeit zwischen Aufgabenstellung und Erwartungshorizont vorliegt [BY, 2016 (3)].¹⁴

In Bezug auf drei weitere Fälle wurden die Aufgabenstellungen aufgrund des in Kapitel 2.1 dargestellten Forschungsfokus' der Studie ausgeschlossen: In zwei Fällen wurden Sachurteile verlangt, die eine objektive Lösung besitzen, so dass das Kriterium der Offenheit nicht erfüllt ist [BY, 2019 (2) und BY, 2019 (4)]. Ein weiterer Fall forderte ein Urteil über die Passung von Bild und Text, so dass religiöse und ethische Kategorien im Urteilsbildungsprozess eine untergeordnete Rolle spielen, weshalb auch dieser Fall aus dem Untersuchungskorpus ausgeschlossen wurde [BY, 2015 (2)].

Auf der anderen Seite sind sechs Aufgabenstellungen in das Untersuchungskorpus aufgenommen worden, die der Definition in Kapitel 5.2 entsprechen, deren Erwartungshorizonte jedoch keine eigenständige Urteilsfähigkeit

¹³ Zu erwähnen ist in diesem Zusammenhang noch die Aufgabenklasse mit den Operatoren »Konsequenzen aufzeigen/Perspektiven entwickeln« (vgl. Kap. 5.2.3 Grenz- und Sonderfälle). In ihnen wird beispielsweise gefordert, dass die Perspektiven »sachkundig und differenziert entwickelt werden« sollen [E-BY, 2015 (1)]. Die Forderung einer Positionierung wird jedoch auch in dieser Aufgabenklasse nicht erhoben.

¹⁴ Vgl. Anhang IV (Urteilsbildung mit Perspektivübernahme).

erwähnen. Besonders auffällig ist dabei die Gruppe der Aufgabenstellungen, die eine Urteilsbildung mit einer Gestaltungsleistung verbinden, also die Operatoren »gestalten« und »entwerfen« nutzen. Die Erwartungshorizonte der Aufgaben geben zwar Hinweise zu den Gestaltungsaspekten wie beispielsweise, dass ein zu entwerfender Zeitungsaufsatz nicht journalistisch ausgefeilt, aber adressaten- und situationsbezogen sein sollte [BY, 2014 (2)]. Sie geben jedoch – mit Ausnahme einer Aufgabe [BY, 2015 (3)] von insgesamt sechs – keinen Hinweis darauf, nach welchen Bewertungskriterien eine Positionierung zu den aufgeworfenen ethischen Fragestellungen bewertet werden soll.

Ebenso nimmt der Erwartungshorizont zur Aufgabenstellung mit dem als »Grenzfall« diskutierten Operator »erörtern«, in welcher die Möglichkeiten und Grenzen der Umsetzung des Pressecodex erörtert werden sollen, keinen Bezug auf eine Positionierung des Prüflings (vgl. Kapitel 5.2.3).

Niedersachsen

In der allgemeinen Benennung der im jeweiligen Aufgabenset zu diagnostizierenden prozessbezogenen Kompetenzen, welche in jedem Erwartungshorizont Niedersachsens vorgenommen wird, ist in durchweg allen Aufgabensets des Untersuchungszeitraums »Urteilskompetenz« als eine zu diagnostizierende Kompetenz genannt. Die Analyse der Erwartungshorizonte aller 24 im Untersuchungszeitraum gestellten Aufgabensets (insgesamt 86 Aufgabenstellungen) hat jedoch ergeben, dass die Forderung einer Stellungnahme oder Positionierung in den expliziten Erwartungsbeschreibungen in vier Aufgabensets nicht noch einmal Erwähnung findet, so dass in diesen vier Fällen eine Diskrepanz besteht zwischen der im Allgemeinen verlangten Kompetenz und des im Konkreten geforderten. Da die Aufgabenstellungen aus Niedersachsen schwerpunktmäßig den drei vorgegebenen Anforderungsbereichen (AFB) zugeordnet werden und da sich Urteilsfähigkeit dem AFB III zuordnen lässt, sind in diesen vier Fällen die Aufgabenstellungen zu AFB III besonders zu betrachten: Es handelt sich in drei Fällen um Aufgabenstellungen mit dem Operator »Perspektiven entwickeln/Konsequenzen aufzeigen« (vgl. Kapitel 5.2.3 Grenz- und Sonderfälle). Die Erwartungshorizonte fordern zwar, dass die Perspektiven »differenziert und eigenständig« entfaltet werden [bspw. NI, 2015, gA (1)], sie fordern jedoch keine eigenständige Urteilsbildung.¹⁵ Ein vierter Fall, in dem Urteilsfähigkeit zwar unter den »prozessbezogenen Kompetenzen« benannt, jedoch in der Aufgabenstellung keine Positionierung gefordert wird, betrifft die ebenfalls bereits im Kapitel »Grenz- und Sonderfälle« diskutierte Aufgabenstellung:

¹⁵ Wie im Kapitel »Grenz- und Sonderfälle« (Kap. 5.2.3) diskutiert, wurden Aufgabenstellungen dieser Klasse auch nicht in das Untersuchungsmaterial aufgenommen.

»4. Überprüfen Sie, ob das Poster der Initiative »Brot für die Welt« (M2) eine geeignete Veranschaulichung der Inhalte des Textes von Helmut Fischer darstellt.« [NI 2014 eA (2)]

Gefordert werden »eigenständige[], begründete[] und differenzierte[] Ergebnisse[]«, jedoch keine Urteilsbildung.

Insgesamt betrachtet wird eine »eigene Stellungnahme« [bspw. NI, 2014, gA (1)], eine »eigene, differenziert begründete Positionierung« [bspw. NI, 2014, eA (1)], eine »differenzierte Beurteilung« [bspw. NI, 2015, gA (2)] oder eine adäquate Forderung in 20 Fällen verlangt. Diese Zahl entspricht der in das Untersuchungsmaterial aufgenommenen Aufgabenzahl. Die in Niedersachsen untersuchten Aufgabenstellungen orientieren sich an den Operatorendefinitionen der EPA, so dass die Zuordnung in der Regel eindeutig zu treffen war.¹⁶

Thüringen

In Thüringen werden die Darstellung »eine[r] begründete[n] Haltung zum Thema« [TH, 2014, eA (2)], eine »subjektiv wertende Auseinandersetzung mit dem Material« [TH, 2015, eA (2)] sowie eine »begründete Positionierung« [TH, 2017, eA (1)] in zwölf Fällen mit gleicher oder ähnlicher Formulierung gefordert. Darunter befindet sich auch – ähnlich wie im oben diskutierten Beispiel aus Bayern – eine Aufgabenstellung, welche eine Antwort »aus christlicher Perspektive« fordert und dann im Erwartungshorizont unter der Überschrift »Entwerfen, Stellung nehmen, Begründen« eine »eigene[] begründete[] Argumentation unter Einbeziehung einer sinnvollen Auswahl der Argumente aus den Materialien, christlichen und sonstigen Wissens« [TH, 2015, eA (1)] fordert, ohne nochmals auf den in der Aufgabenstellung geforderten Perspektivwechsel einzugehen.¹⁷

Darüber hinaus befindet sich ähnlich wie in Bayern unter den Aufgabenstellungen eine Fragestellung mit dem Operator »entwerfen«, in welcher im Erwartungshorizont die Urteilsbildung unangesprochen bleibt, die jedoch aufgrund der in Kapitel 5.2.2 gegebenen Definition in das Untersuchungsmaterial aufgenommen wurde. Es handelt sich um folgende Aufgabenstellung:

»1. Entwerfen Sie in Auseinandersetzung mit den Worten von Margot Käßmann eine Rede zur Abschlussfeier der Absolventen Ihres Jahrgangs. [...]« [TH, 2014, eA (1)].

¹⁶ Dies wird auch anhand der im Kapitel »Grenz- und Sonderfälle« besprochenen Fälle deutlich, welche überwiegend aus Bayern und Thüringen stammen (vgl. Kap. 5.2.3).

¹⁷ Aufgrund des geforderten Perspektivwechsels ist diese Aufgabenstellung nicht mit im Datenkorpus enthalten.

Anders als in Bayern, ist es jedoch nicht generell der Fall, dass Aufgabenstellungen mit den Operatoren »gestalten« und »entwerfen« Urteilsbildung im Erwartungshorizont ausblenden. In mehreren Fällen werden Stellungnahmen unter der Überschrift »Entwerfen und Gestalten« in den Erwartungshorizonten gefordert.

Eine dritte Differenz zwischen den in das Untersuchungskorpus aufgenommenen Aufgaben und den Erwartungshorizonten betrifft eine Aufgabe aus dem Jahr 2019, die lautet:

»Entwerfen Sie, ausgehend von der Rede Martin Luther Kings jr., ein Essay für die Schülerzeitung, indem Sie – unter Nutzung der christlichen Anthropologie und der christlichen Ethik – die Forderungen Luther Kings auf die heutige Zeit anwenden. Analysieren Sie dazu die Redeauszüge (Material) und benennen Sie die wichtigsten Forderungen, die, im Sinne des kategorischen Imperativs, bleibend, [sic] Gültigkeit haben sollten.« [TH, 2019, eA (2)]

Die Aufgabe verwendet die Operatoren »entwerfen«, »anwenden«, »analysieren« und »benennen«. Der Erwartungshorizont listet unter dem Operator »entwerfen« die Forderung »eigene Positionierung« mit auf. Irritierend ist dabei jedoch die Überschrift: »Beispielhaft können folgende Argumente verwandt werden«. Statt Argumente folgen »Qualitätsmerkmale« wie »Schülerzeitungssessay – Leserbezug«, »angemessener sprachlicher Ausdruck«; Stichpunkte wie »Bezug zu Martin Luther King jr.« und eben die »eigene Positionierung«. Ob die eigene Positionierung fakultativ ist, wie das »Beispielhaft« der Überschrift vermuten lässt, bleibt unklar. Die Aufgabenstellung wurde aufgrund ihrer Formulierung, die Forderungen MARTIN LUTHER KINGS sollen auf die heutige Zeit *angewendet* werden, nicht mit in das Untersuchungskorpus aufgenommen, da sie keine eindeutige Aufforderung zur Urteilsbildung enthält.

Vergleichende Perspektiven

Eigene Positionierungen werden in den Erwartungshorizonten der drei untersuchten Bundesländer unterschiedlich oft gefordert. In Niedersachsen werden mit 20 Erwähnungen doppelt so häufig Urteilsbildungen in den Erwartungshorizonten gefordert, wie es in den Erwartungshorizonten in Bayern der Fall ist, während von beiden Bundesländern die gleiche Anzahl an Aufgabensets in die Untersuchung eingeflossen sind (24). Bei diesem Vergleich ist jedoch zu beachten, dass die Anzahl der Aufgabenstellungen im Untersuchungszeitraum in Bayern mit 166 Aufgaben um 80 Aufgaben höher war als in Niedersachsen (86 Aufgabenstellungen im Untersuchungszeitraum). Im niedersächsischen Abitur werden also auf weniger Aufgabenstellungen mehr eigenständige Urteilsbildungen der Prüflinge gefordert. Thüringen erwartet in zwölf Aufgabenstellungen innerhalb der sechs untersuchten Jahre von den Prüflingen eine eigenständige

Urteilsbildung. Da jedoch in diesem Zeitraum in Thüringen nur 12 Aufgabensets mit insgesamt 38 Aufgabenstellungen gestellt wurden, ist die relative Häufigkeit von Urteilsbildungsaufgaben höher als in den anderen Bundesländern, wie folgende Tabelle zeigt:

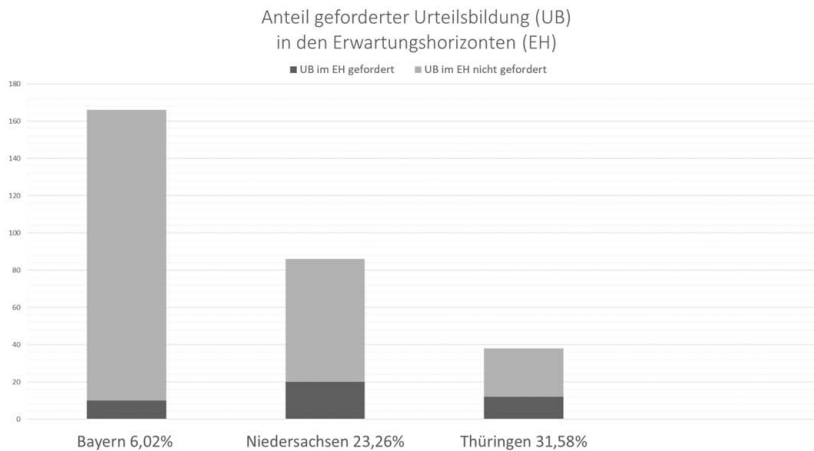


Abb. 17: Säulendiagramm zur relativen Häufigkeit von Urteilsbildungen im schriftlichen Abitur in den Bundesländern Bayern, Niedersachsen und Thüringen.

6.3.2 Qualitätsmerkmale von Urteilsbildung in den Erwartungshorizonten (F)

Werfen wir einen Blick auf den Inhalt der Erwartungshorizonte: In Kategorie F wurden die Erwartungshorizonte der 42 Fälle hinsichtlich der in ihnen genannten Qualitätsmerkmale untersucht.¹⁸ In der Darstellung der Ergebnisse sind die Qualitätsmerkmale nach der Häufigkeit ihrer Nennung geordnet. Die in Klammern befindliche Zahl gibt an, in wie vielen Lösungsskizzen der Teilaufgaben das Merkmal genannt wurde – in Bezug auf ein und dieselbe Teilaufgabenstellung wiederholt genannte Qualitätsmerkmale sind nicht doppelt in die Zählung eingegangen.¹⁹ Differenziert wird zwischen allgemeinen Qualitätsmerkmalen, die sich auf die Urteilsbildung beziehen, und aufgabentextspezifischen Qualitätsmerkmalen, welche dezidiert eine in der Aufgabe formulierte

¹⁸ Allgemeingültige Aufzählungen von Qualitätsmerkmalen wurden nicht berücksichtigt, sondern ausschließlich die zu Aufgaben von Urteilsfähigkeit genannten.

¹⁹ Weil in Thüringen keine direkte Zuordnung der Anforderungen zu den Teilaufgaben besteht und in einem der Aufgabensets zwei Teilaufgaben zur Urteilsbildung vorkommen, werden die Angaben als für beide Teilaufgaben gültig doppelt in die Zählung aufgenommen, um das Verhältnis zur Gesamtaufgabenanzahl nicht zu verfälschen.

Forderung aufgreifen und einholen (vgl. Kapitel 5.3.3). So taucht beispielsweise die Multiperspektivität in Bayern sowohl unter der Kategorie »aufgabentext-unspezifische«, als auch in der Kategorie »aufgabentextspezifische Qualitätsmerkmale« auf, da sie in einem Fall konkret in der Aufgabenstellung verlangt wird und in zwei anderen Fällen als ein allgemeines Qualitätsmerkmal der Urteilsbildung erscheint, auf das in der Aufgabenstellung nicht noch einmal deziert hingewiesen wird.

Am häufigsten wird in den Erwartungshorizonten die Positionierung an sich als ein Qualitätsmerkmal genannt (in 35 der 42 untersuchten Fälle). Dass eine Begründung vorhanden sein soll, wird in mehr als der Hälfte der Fälle als Qualitätsmerkmal genannt. Ein Aufgabenbezug wird in knapp der Hälfte der Lösungsskizzen erwartet. Darüber hinaus gibt es wenige Qualitätsmerkmale, die wiederholt vorkommen.

Folgende, nicht in der Aufgabenformulierung genannte Qualitätsmerkmale sind in mindestens zwei Bundesländern mehr als einmal genannt worden (n = 42):

- Stellungnahme/Antwort/Urteil (35)
- Begründung (26)
- Aufgabenbezug (20)
- Differenziertheit²⁰ (16)
- Eigenständigkeit (15)
- Fundierte Kenntnisse/sachgerecht (14)
- Schlüssigkeit (6)
- Verständlichkeit/sprachliche Gestaltung (6)²¹

Darüber hinaus unterscheiden sich die genannten Kriterien auch hinsichtlich ihres naheliegenden Bewertungsmodus: Die drei meistgenannten Kriterien »Stellungnahme«, »Begründung« und »Aufgabenbezug« legen einen binären Modus nahe, in welchem eine Zuordnung zu den Polen »vorhanden«/»nicht-vorhanden« stattfindet (vgl. Kapitel 3.3.2).²² Sie sind Qualitätsmerkmale, da ein Urteilsentscheid durch ihr Vorhandensein an Qualität gewinnt, jedoch sind Er-

²⁰ Die Differenziertheit wird häufig in Bezug auf die Begründung geforderte (bspw. eine »überzeugende und differenzierte Begründung«).

²¹ In Niedersachsen taucht in jedem Erwartungshorizont diese oder ähnliche allgemeine Formulierung auf: »Insgesamt sind die Ausführungen deutlich auf die Aufgabenstellung bezogen sowie sprachlich und gedanklich präzise«. Da sich diese Aussage jedoch nicht spezifisch auf Aufgaben der Urteilsbildung bezieht, wird sie nicht in diese Aufstellung aufgenommen. Der sprachlichen Gestaltung und dem Merkmal »Aufgabenbezug« würden sonst mehr Gewicht zukommen. In Thüringen tauchen »Aufgabenbezug« und »sprachliche Gestaltung« ebenso in einer Aufzählung allgemeiner Qualitätsmerkmale auf. Sie sind jedoch dort in Aufgabenstellungen zur Urteilsfähigkeit mitunter nochmals genannt.

²² Vgl. SCHÖNEMANN/THÜNEMANN/ZÜLSDORF-KERSTING, Was können Abiturienten?, 34.

gänzungen nötig, um sie graduell messbar zu machen. In den Erwartungshorizonten häufig genannte Qualitätsmerkmale, welche einen graduellen Modus nahelegen, sind »Differenziertheit«, »Eigenständigkeit« und »fundierte Kenntnisse/sachgerecht«. Da jedoch keine Ausprägungsgrade genannt werden und es somit keinen Maßstab dafür gibt, wann eine Begründung auf welchem Niveaugarad als »differenzierte Begründung« bewertet werden kann, muss die korrigierende Lehrkraft diese Maßstäbe selbst festlegen, was zu einem Verlust an Auswertungsobjektivität führt. Ebenso verhält es sich mit den weniger als halb so oft genannten Qualitätsmerkmalen »Schlüssigkeit« und »Verständlichkeit/sprachliche Gestaltung«: Auch sie legen einen graduellen Modus nahe, welcher jedoch einer weiteren Differenzierung der Ausprägungsgrade bedürfte, um einen vergleichbaren Maßstab zu haben.

In der Formulierung der Aufgabe geforderte Merkmale, die in mindestens zwei Bundesländern mehr als einmal genannt wurden, richten sich vor allem auf in den Aufgabenstellungen geforderte herzustellende Bezüge (n = 42):

- Materialbezug (21)
- Einbezug geforderter Reflexionskriterien (z.B. christliche Anthropologie, Bibelstellen, persönliche Erfahrungen etc.) (15)
- Adressatenbezug (13)
- Situationsbezug (11)

Im Folgenden werden die Qualitätsmerkmale der untersuchten Bundesländer einzeln aufgeführt, um die Breite der genannten Qualitätsmerkmale darzustellen.

Bayern

Aufgabentextunspezifische Qualitätsmerkmale Bayern (n = 11)

- Stellungnahme/Antwort/Urteil (5)
 - Differenziertheit (4)
 - Begründung (4)
 - Kritisch (3)
 - Verständliche/angemessene Sprache (3)
 - Multiperspektivität²³ (2)
 - Schlüssigkeit (2)
 - Eigenständigkeit (2)
 - Reflektiertheit (1)
 - Überzeugungskraft (1)
 - Ggf. Beispiele einbringen (1)
-

²³ Multiperspektivität erfasst die Forderungen, sowohl meinungskonforme als auch non-konforme Argumente zu betrachten, wie beispielsweise die Forderung, »[s]owohl Chancen als auch Grenzen« differenziert zu würdigen [BY, 2014 (2)].

-
- Strukturierte Darstellung (1)
-

Aufgabentextspezifische Qualitätsmerkmale Bayern

- Einbezug der geforderten Reflexionskriterien²⁴ (8)
 - Adressatenbezug (6)
 - Situationsbezug (6)
 - Materialbezug (5)
 - Multiperspektivität (1)
 - Kritische Würdigung der zu diskutierenden Position (1)
 - Fragestellung zunächst verifizieren²⁵ (1)
-

Explizit nivellierte Qualitätskriterien Bayern

- Journalistische/rhetorische Ausgefeiltheit (2)
-

Tab. 21: Qualitätsmerkmale der Urteilsbildung aus den Erwartungshorizonten bayerischer Abituraufgabenstellungen in den Jahren 2014–2019.

Niedersachsen²⁶

Aufgabentextunspezifische Qualitätsmerkmale Niedersachsen (n = 20)

- Stellungnahme/Standpunkt/Urteil/Schlussfolgerung/Positionierung etc. (20)
 - Eigenständigkeit²⁷ (13)
 - Begründung (13)
 - Differenziertheit (12)
 - Fundierte Kenntnisse/inhaltlich fundiert/fachgemäße Kenntnisse/»sachgemäß« (9)
 - Durchdringungstiefe (8)
 - Plausibilität (7)
 - (Angemessen)/(inhaltlich) aspektreich (5)
 - Schlüssigkeit der Verknüpfungen der Aspekte (4)
 - Nachvollziehbarkeit (2)
-

²⁴ In diese Kategorie sind Reflexionsforderungen, wie bspw. die Berücksichtigung von christlicher Anthropologie und christlichen Zukunftshoffnung aufgenommen worden.

²⁵ In diesem Fall sollten laut Erwartungshorizont zunächst grundlegende Aussagen des Apostolikums benannt und als Antworten auf Fragen des Lebens kenntlich gemacht werden [BY, 2019 (4)]. Die Fragestellung lautet, ob das apostolische Glaubensbekenntnis umfassende Antworten auf die »letzten Fragen« des Lebens geben kann.

²⁶ Die beiden jeweils nahezu identischen Erwartungshorizonte [NI, 2017, gA (1)] und [NI, 2017, eA (2)] sowie [NI, 2019, gA (1)] und [NI, 2019, eA (1)] (vgl. Kap. 7.3.4) wurden unabhängig voneinander in die Aufstellung aufgenommen, um keine Verzerrungen zu erzeugen.

²⁷ Diese Kategorie erfasst neben dem Adjektiv »eigenständig« auch Formulierungen, in denen das Adjektiv »eigen« den Forderungen einer Stellungnahme oder einem Urteil vorangestellt ist (»eine eigene Stellungnahme«).

-
- Abwägung/abwägende Prüfung (2)
 - Verdeutlichung des jeweiligen ethischen Grundprinzips, von dem aus geurteilt wird (1)
 - Angemessenes Textverständnis (1)
 - Ausführlichkeit (1)
-

Aufgabentextspezifische Qualitätsmerkmale Niedersachsen

- Materialbezug (9)
 - Briefform (3)
 - Kommentarform (2)
 - Adressatenbezug (2)
 - Rhetorische Mittel (2, davon 1x »ggf.«)
 - Plausible Kriterien (2)
 - Kritische Würdigung/Überprüfung der geg. Position (2)
 - Kreativität (1)
 - Bezug zu mind. einer Leiderfahrung (1)
 - Überzeugende und einfallsreiche Gestaltung des Briefes (1)
 - Möglich: Verweis auf bibl. Erzählungen (1)
 - Texterläuterung (1)
 - Religionskritische Positionen identifizieren (1)
-

Explizit relativierte Qualitätskriterien Niedersachsen

- Quantitativer Umfang der Argumente²⁸ (10)
-

Tab. 22: Qualitätsmerkmale der Urteilsbildung aus den Erwartungshorizonten niedersächsischer Abituraufgabenstellungen in den Jahren 2014–2019.

Thüringen

Aufgabentextunspezifische Qualitätsmerkmale Thüringen (n = 11)

- Haltung/subjektiv wertende Auseinandersetzung/ Positionierung/Stellungnahme (10)
 - Begründung (9)
 - Aufgabenbezug (6)
 - Sachgerecht (5)
 - Sachliche, klare und präzise Formulierung/angemessene sprachliche Gestaltung/ Verständlichkeit (3)
 - Reflektiertheit (2)
 - Ausführlichkeit der Argumente (1)
-

²⁸ Ein Beispiel für eine solche Relativierung der Quantität bietet folgende Formulierung: »Wichtiger als der quantitative Umfang thematisierter Aspekte sind ein deutlich erkennbarer Bezug der Ausführungen auf das Titelblatt des »SPIEGEL«, die Berücksichtigung ausgewählter Textaussagen, sowie die Differenziertheit und die Plausibilität der Darlegung mit Blick auf die eigene Positionierung« [E-NI, 2014, eA (1)].

-
- Schlüssigkeit (1)
 - Bezugnahme auf selbst gewählte Beispiele (1)
 - Kohärenter Aufbau (1)
 - Multiperspektivität (1)
 - Erschließen von Sinnzusammenhängen (1)
 - Verwendung von Fachwissen (1)
 - Verwendung von Fachmethoden (1)
 - Fachtermini (1)
 - Eigenständigkeit (1)
-

Aufgabentextspezifische Qualitätsmerkmale Thüringen

- Einbezug der Materialien/Schlussfolgerungen aus Materialien (7; davon enthalten 4 biblische Texte)
 - Einbezug biblischer Textstellen, ohne dass sie im Material gegeben sind (2)
 - Kreativität (7)
 - Adressatenbezug/Zielgruppenorientierung (5)
 - Situationsgerecht (5)
 - Einbezug persönlicher Erfahrungen (5)
 - In pointierter Form Denkanstöße schaffen (2)
 - Offene Schreibform (2)
 - In expressiver und appellativer Form Meinungsbildung beeinflussen (1)
 - erkennbarer Gegenwartsbezug (1)
 - Redeaufbau, Gedankenführung, rhetorische Strategien und sprachliche Gestaltungsweise von Trauerreden (1)
 - mindestens 2 Beispiele (pro oder contra) nach eigener Wahl (1)
 - Sachverhalt im logischen Zusammenhang unter Verwendung der Fachsprache wiedergeben (1)
 - Briefform (Anrede, Hauptteil, Unterzeichnung) (1)
-

Explizit nivellierte/relativierte Qualitätskriterien Thüringen

- Keine
-

Tab. 23: Qualitätsmerkmale der Urteilsbildung aus den Erwartungshorizonten thüringischer Abituraufgabenstellungen in den Jahren 2014–2019.

Die als aufgabentextspezifisch benannten Merkmale beziehen sich überwiegend auf Gestaltungsaspekte und Textformen. Sie zählen in allen drei Bundesländern zu den Bewertungskriterien für Urteilsbildungsaufgaben. Interessant ist in diesem Zusammenhang, dass die »journalistische« bzw. »rhetorische Ausgefeiltheit« in Bayern zweimal explizit als Qualitätsmerkmal nivelliert wird. Dennoch sind auch dort formale Kriterien wie der Situations-, Material- und Adressatenbezug Merkmale, die in die Bewertung einfließen sollen.

In den Erwartungshorizonten aus Niedersachsen wird die Bedeutung des quantitativen Umfangs von Argumenten zehnmal als zweitrangig relativiert. Dennoch wird auch fünfmal »aspektreich« als ein Qualitätsmerkmal benannt.

6.3.3 Explizite Hinweise zum Umgang mit Schülermeinungen (G)

Bei der Durchsicht der Erwartungshorizonte ist aufgefallen, dass in einigen der Umgang mit Schülermeinungen explizit angesprochen wird, in anderen nicht. Um zu prüfen, wie häufig dies im Untersuchungszeitraum der Fall ist, wurde induktiv eine entsprechende Kategorie gebildet. In 16 der 42 Urteilsbildungsaufgaben wird die Offenheit der Urteilsbildung explizit thematisiert. Es sind dabei erhebliche Unterschiede zwischen den Bundesländern auszumachen: Während entsprechende Stellen in den aus Niedersachsen stammenden Erwartungshorizonten häufig zu finden sind, kommen sie in den Erwartungshorizonten aus Bayern selten und in thüringischen Erwartungshorizonten gar nicht vor.

In Bayern wird in zwei der elf analysierten Aufgabenstellungen im Erwartungshorizont explizit darauf verwiesen, dass die Urteilsfindung hinsichtlich der Positionierung offen ist. Beide Erwartungshorizonte stammen aus dem Jahr 2015 [Erwartungshorizont BY, 2015 (3)] und [Erwartungshorizont BY, 2015 (4)]. So heißt es beispielsweise:

»Es steht dem Prüfling frei, sich für oder gegen die Aufnahme von Demut in die Schulverfassung zu entscheiden; [...]« [E-BY, 2015 (3)].

In den niedersächsischen Erwartungshorizonten werden in vierzehn der zwanzig analysierten Aufgabenstellungen explizit darauf eingegangen, dass verschiedene Positionierungen möglich sind. In mehreren Fällen werden diese sogar genannt. Zwei Beispiele sollen dies illustrieren. Ein Satz, der in gleicher und ähnlicher Weise wiederholt in den Erwartungshorizonten steht, ist folgender:

»Aufgrund der eigenen Position des Prüflings und des vorangegangenen Unterrichts sind verschiedene Bearbeitungen dieser Aufgabe möglich« [E-NI, 2015, gA (2)].

Darüber hinaus tauchen Aussagen auf, die die Möglichkeit der Zustimmung oder Ablehnung explizit erwähnen, wie exemplarisch an folgender Äußerung gezeigt werden kann:

»Die Prüflinge müssen sich in einem Brief zu diesen Forderungen verhalten und dabei begründen, ob sie ihnen zustimmen, sie als zu weitgehend oder als zu schwach bewerten würden« [E-NI, 2015, eA (2)].

In den Aufgabenstellungen aus Thüringen lassen sich explizite Hinweise dieser Art nicht finden. In zwei der elf analysierten Aufgabenstellungen findet sich ein Hinweis darauf, dass es sich um subjektive Auseinandersetzungen mit der Fragestellung handeln soll. Wie jedoch mit der Subjektivität umzugehen ist, findet keine Erwähnung.

6.3.4 Formulierung konträrer Positionen in den Erwartungshorizonten (H)

Eng mit der vorangegangenen Kategorie verknüpft, ist die Analysefrage, inwieweit konträre Positionen zur Fragestellung der Urteilsbildung in den Erwartungshorizonten inhaltlich ausformuliert werden. Dies ist in 19 der 42 Erwartungshorizonte der Fall. Die Verteilung dieser Fälle auf die Bundesländer ergibt ein ähnliches Bild wie in der vorangegangenen Kategorie: In den Erwartungshorizonten aus Bayern sind in nur drei der elf Aufgabenstellungen konträre Positionen zur aufgeworfenen Fragestellung aufgeführt (~27%). In aus Niedersachsen stammenden Erwartungshorizonten werden in 16 von 20 analysierten Fällen konträre Positionen zur Fragestellung benannt (80%). Keiner der Erwartungshorizonte der elf aus Thüringen analysierten Aufgabenstellungen gibt zwei mögliche Positionen hinsichtlich der Fragestellung an (0%).

6.4 Zusammenfassung der kategorienorientierten Ergebnisse

In 60% aller Aufgabensets wurde mindestens eine Aufgabenstellung gestellt, die eine eigenständige Urteilsbildung in kontroversen, ethischen und/oder religiösen Fragen forderte, so dass man sagen kann, dass Urteilsbildung zwar regelmäßig geprüft wird, jedoch auch nicht zwangsläufig in Abitursets des Faches Evangelische Religionslehre vorkommt. Urteilsbildungsaufgaben werden in der Regel in Zusammenhang mit einem der Aufgabe zugrunde liegendem Material gestellt und umfassen sowohl Text- als auch Gestaltungsaufgaben. Darüber hinaus ist der Befund zentral, dass Urteilsbildungsanforderungen, die in Zusammenhang mit Gestaltungsaufgaben gestellt werden, in den Erwartungshorizonten gar nicht oder nur unzureichend aufgenommen sind, was vor allem in den Erwartungshorizonten aus Bayern besonders deutlich wird: In diesen Erwartungshorizonten fehlen Hinweise über Bewertungskriterien und Qualitätsmerkmale der geforderten Urteilsbildung, da der Fokus auf den Gestaltungsaspekten der Aufgabenstellung liegt.

Sowohl unter den Aufgabensets aus Bayern als auch unter denen aus Thüringen ist je eine Aufgabe zu finden, in welcher die Aufgabenformulierung einen

Perspektivwechsel im Zusammenhang mit der Urteilsbildung fordert, welcher dann im Erwartungshorizont nicht mehr erwähnt wird. Stattdessen wird im Erwartungshorizont eine eigenständige Positionierung der Prüflinge verlangt, die nicht aus der Aufgabenstellung zu erschließen ist. Eine weitere Unstimmigkeit zeigt sich in einer Aufgabenstellung, die zwar auf Aufgabenebene keine Urteilsbildung fordert, jedoch im Erwartungshorizont [TH, 2019, eA (2)].

Zu den in den Erwartungshorizonten genannten Qualitätsmerkmalen lässt sich feststellen, dass sich die Übereinstimmungen der genannten Qualitätsmerkmale zwischen den 42 Fällen der Studie sowie zwischen den drei untersuchten Bundesländern auf wenige Qualitätsmerkmale beschränken, welche ihrerseits sehr allgemein gehalten sind und keine graduelle, d.h. hierarchisch nach Niveaustufen differenzierende Bewertung zulassen. Es lassen sich acht Qualitätsmerkmale benennen, welche in mindestens zwei Bundesländern mehr als einmal genannt werden.



Abb. 18: Qualitätsmerkmale für Urteilsbildung, die in mindestens zwei der untersuchten Bundesländer mehr als einmal genannt werden. Der in Klammern angegebene Wert gibt die Häufigkeit der Nennung an, wobei $n = 42$ ist.

Neben diesen acht übereinstimmenden Qualitätsmerkmalen werden weitere Qualitätsmerkmale in den Erwartungshorizonten genannt. Ein der Mehrheit der Aufgabenstellungen zugrunde liegendes einheitliches Konstrukt der Kompetenz Urteilsfähigkeit lässt sich jedoch nicht nachweisen. Interessant ist darüber hinaus, dass allein in den aus Niedersachsen stammenden Erwartungshorizonten wiederholt quantitative Kriterien zur Bewertung der Urteilsbildung, wie beispielsweise den Umfang der Argumente, nivelliert werden.

Hinsichtlich der Themen der Fragestellungen sowie der geforderten Reflektionskriterien sind bundeslandspezifische Konventionen zu beobachten: Während im bayerischen Abitur in erster Linie ethische Fragestellungen ohne dezidiert religiösen Bezug aufgeworfen werden und die christliche Weltsicht als ein

Reflexionskriterium der eigenen Urteilsbildung zum Tragen kommt, wird in den Aufgabenstellungen aus Thüringen der Bezug zur Religion in erster Linie über die Fragestellung hergestellt. In niedersächsischen Abiturprüfungen hingegen soll über konkrete, in der Regel christliche Positionierungen geurteilt werden, welche sich sowohl auf Fragestellungen mit dezidiert religiösen Bezügen als auch auf Fragestellungen ohne dezidiert religiöse Bezüge richten. Die zu beurteilenden Positionierungen werden über die den Aufgabenstellungen zugrunde liegenden (meist theologischen) Sachtexte vermittelt. Den bayerischen Aufgabenstellungen liegen hingegen vorzugsweise Zeitungsartikel und -interviews zugrunde, während im thüringischen Abitur vermehrt auch lyrische Materialien wie Liedtexte, Gedichte und persönlich verfasste Glaubensbekenntnisse zum Einsatz kommen, die wiederum in den Fällen aus Niedersachsen und Bayern nicht zu finden sind. Darüber hinaus ist im thüringischen Abitur der Einbezug von Bibelstellen am häufigsten gefordert und Bibelverse und -perikopen werden als Materialgrundlage einbezogen, während Bibelverse als Primärquelle in den Materialien der Aufgabenstellungen aus Niedersachsen und Bayern keine Rolle spielen. Sie kommen dort ausschließlich in kontextualisierter Form – beispielsweise in den theologischen Sachtexten – vor.

Große Unterschiede bestehen zwischen den Erwartungshorizonten der drei Bundesländer hinsichtlich einer expliziten Thematisierung der Offenheit der Positionierungen auf einer Metaebene sowie der Nennung konträrer Positionen zur Fragestellung. Während diese Aspekte in niedersächsischen Erwartungshorizonten häufig vorkommen, sind sie in Erwartungshorizonten aus Bayern selten. In Erwartungshorizonten aus Thüringen kommen sie hingegen gar nicht vor.

Die kategorienbasierte Analyse zeigt somit, dass die Prüfungskonventionen und die Bewertungsvorgaben von Urteilsfähigkeit länderspezifisch variieren. Die zentralen Charakteristika der drei Bundesländer münden gemeinsam mit den Ergebnissen der im Folgenden dargestellten vertiefenden Einzelfallinterpretationen in Profile der Bundesländer (vgl. Kapitel 8).

7 Fallorientierte Ergebnisse

7.1 Methodische Einführung

Wie bereits in Kapitel 5.4.3 zum Analysevorgehen und zur Methodenreflexion geschildert, soll die kategorienorientierte Auswertung durch eine vertiefende Einzelfallinterpretation ergänzt werden, um die kontextuelle Einbettung einzelner Fälle deutlich werden zu lassen, Ergebnisse der kategorienorientierten Analyse nochmals zu illustrieren und theoretische Schlussfolgerungen ziehen zu können.¹ Während in der Kategorienanalyse Aufgabenstellungen und Erwartungshorizonte getrennt voneinander analysiert wurden, wird in der vertiefenden Einzelfallanalyse beides im Kontext des gesamten Aufgabensets gemeinsam betrachtet. Wie bereits im Namen der Methodik ersichtlich, handelt es sich um eine Interpretation der Einzelfälle, für die nach KUCKARTZ »nicht solch strikte Regeln wie für die tabellarischen Fallübersichten und die vorangegangenen Analysephasen« gelten.² Es kommen, wie von KUCKARTZ vorgeschlagen, hermeneutische Techniken zum Einsatz, welche auch zu evaluativen Aussagen führen.³

Die Auswahl der Fälle erfolgte nach Besonderheiten, die während der kategorienorientierten Analyse wahrgenommen und in den Fallmatrizen vermerkt wurden.⁴ Leitend für die Einzelfallanalyse ist eine konkrete Fragestellung, unter welcher die Fälle betrachtet werden.⁵ Im Zentrum der hiesigen Fallanalysen steht die Fragestellung, welchen Maßstabsgehalt die in den Erwartungshorizonten gegebenen Qualitätsmerkmale aufweisen und in welchem Maße durch die Erwartungshorizonte Auswertungsobjektivität sichergestellt werden kann. Be-

¹ Vgl. KUCKARTZ, Qualitative Inhaltsanalyse, 116 f.

² A. a. O., 117.

³ Vgl. ebd.

⁴ Vgl. SCHMIDT, Analyse von Leitfadeninterviews, 482–484.

⁵ Vgl. a. a. O., 482.

reits die Fragestellung zeigt die evaluative Zielrichtung der Einzelfallinterpretationen, da sie eine Bewertung der Fälle deutlich intendiert.

Die Fallanalyse basiert auf einem kompetenzorientierten Bildungsverständnis, welches Bildung als die Verknüpfung einer Handlung mit einem Inhalt definiert, bei welcher man eine Haltung zu diesem aufbaut (vgl. Kapitel 2.2): Die Analyse wird daher durch die drei Bereiche *Fähigkeiten und Fertigkeiten der Handlung*, *Wissen über den Gegenstand* sowie *Einstellungen und Haltungen zum Gegenstand* strukturiert. In Bezug auf Urteilsfähigkeit ergeben sich folgende Leitfragen für die Bereiche:

Wissen: Welche zur Bearbeitung der Aufgabenstellung notwendigen Wissensaspekte werden vom Erwartungshorizont benannt? Welche inhaltlichen Bereiche werden bereits durch vorangehende Aufgabenstellungen betrachtet? Werden mögliche Verständnishürden oder Fehler benannt?

Fähigkeiten und Fertigkeiten: Welche zur Bearbeitung der Aufgabenstellung notwendigen Fähigkeiten und Fertigkeiten werden im Erwartungshorizont benannt? Welche Aussagen werden in Bezug auf Qualitätsmerkmale für das Urteilen getroffen?

Einstellungen und Haltung: Thematisieren die Erwartungshorizonte explizit oder implizit die Rolle der Einstellungen und Haltung zum Urteilsgegenstand (beispielsweise durch die Nennung verschiedener Positionen)?

Mit diesen Leitfragen verbinden sich bereits normative Setzungen, da sie implizit mit dem Anspruch verbunden sind, dass der Erwartungshorizont jeder Teilaufgabe alle drei benannten Bereiche abdeckt, mögliche Verständnishürden sowie Qualitätsmerkmale für die geforderte Handlung benennt und die Rolle der Einstellungen und Haltungen der Prüflinge thematisiert. Diese normativen Setzungen lassen sich zum einen aus den Einheitlichen Prüfungsanforderungen (EPA) und zum anderen aus einem kompetenzorientierten Bildungsverständnis ableiten und sollen in erster Linie als heuristisches Modell der Strukturierung dienen, wenngleich niemand in Frage stellen wird, dass die Bereiche eng miteinander verknüpft sind und in den Erwartungshorizonten nicht getrennt voneinander vorliegen.

Zur Einordnung der betrachteten Fälle erscheint auch die Bearbeitungszeit von Interesse. Um einen Richtwert zu generieren, wie viel Zeit einem Prüfling für die geforderte Urteilsbildung zur Verfügung steht, wird der prozentuale Anteil der Zeit entsprechend der prozentualen Gewichtung der Aufgabe errechnet, die durch die Vergabe der Bewertungseinheiten gesetzt ist.⁶

⁶ Dieser Ableitung liegt die Annahme zugrunde, dass die Punktevergabe auch den Prüflingen dazu dient, Schwerpunktsetzungen vorzunehmen und die Gesamtbearbeitungszeit sinnvoll einzuteilen, wenngleich eine flexible Gestaltung der Zeit jederzeit möglich ist. Allerdings sind Urteilsbildungsaufgaben, da sie dem höchsten AFB angehören und Aufgabensets häufig einem schwierigkeitssteigernden Aufbau folgen, meist am Ende des Aufga-

Eingeleitet werden die vertiefenden Fallinterpretationen durch einen Vergleich der formalen Aspekte der Erwartungshorizonte aus den Bundesländern Bayern, Niedersachsen und Thüringen. Vor dem Hintergrund dieses Vergleichs lassen sich die vertiefenden Einzelfallanalysen leichter verstehen und einordnen, während die Fallanalysen ihrerseits den Vergleich exemplarisch illustrieren.

7.2 Formale Ausgestaltung der Erwartungshorizonte im Vergleich zwischen den untersuchten Bundesländern

7.2.1 Formale Aspekte bayerischer Erwartungshorizonte

Die »Korrekturhinweise« aus Bayern leiten mit dem Hinweis ein, dass in selbigen keine vollständigen Lösungen, sondern lediglich »kurze Angaben zur erwarteten Schülerleistung« gegeben werden und dass nicht genannte, gleichwertige Lösungswege gleichberechtigt sind. Sie enthalten Lösungsskizzen zu jeder Aufgabe, welche aus kursivgedruckten Textteilen bestehen, die allgemeine Hinweise und Qualitätsmerkmale zur Bewertung der entsprechenden Aufgabenstellung ausweisen, sowie aus nicht kursivgedruckten Textteilen, welche inhaltliche Aspekte der Lösung benennen. Es kommt mitunter – gerade bei sehr offen formulierten Aufgabenstellungen – vor, dass ausschließlich allgemeine Aspekte zur Lösung benannt werden, ohne dass diese durch konkrete Inhalte ausgeführt sind. Überwiegend sind die Erwartungshorizonte in Bayern in ganzen Sätzen formuliert. Nur selten werden einzelne Aspekte in Stichpunkten genannt.

7.2.2 Formale Aspekte niedersächsischer Erwartungshorizonte

Die aus Niedersachsen stammenden Erwartungshorizonte sind im Vergleich zu den aus Bayern und Thüringen stammenden Erwartungshorizonten die umfangreichsten hinsichtlich der in ihnen enthaltenen Bestandteile: Eingangs werden sowohl die im niedersächsischen Lehrplan vorgegebenen inhaltsbezogenen Kompetenzbereiche (Mensch, Gott, Jesus Christus, Ethik etc.) benannt, auf die das entsprechende Aufgabenset Bezug nimmt, als auch die ebenso im Lehrplan festgeschriebenen und mit dem Kompetenzmodell der EPA übereinstimmenden prozessbezogenen Kompetenzen (Wahrnehmungs- und Darstel-

bensets positioniert (in Niedersachsen ist dies ohne Ausnahme der Fall, in Bayern und Thüringen ist eine solche Steigerung in Tendenzen zu beobachten), so dass es eines konsequenten Zeitmanagements bedarf, um am Ende noch die gesamte für die Aufgabe zur Verfügung stehende Bearbeitungszeit zu haben.

lungskompetenz, Deutungskompetenz etc.), die im entsprechenden Aufgabenset diagnostiziert werden sollen.⁷

Darüber hinaus findet in den Ausführungen zu den einzelnen Aufgabenstellungen jeweils eine schwerpunktmäßige Zuordnung zu einem der drei in den EPA vorgegebenen Anforderungsbereichen (AFB) statt, welche in Bayern nicht und in Thüringen nicht in dieser Form (siehe unten) zu finden ist. Dabei wird im grundlegenden Anforderungsniveau je eine Aufgabenstellung zu jedem AFB stets mit steigendem Komplexitätsgrad formuliert. Im erhöhten Anforderungsniveau entfallen auf AFB I und III jeweils eine Aufgabenstellung und auf AFB II zwei Aufgabenstellungen pro Aufgabenset, wobei sich auch hier die Reihenfolge am steigenden AFB orientiert. Die Anzahl der Aufgaben je Aufgabenset ist aufgrund dieser klaren Zuteilung in Niedersachsen weitestgehend konstant, während sie in den anderen beiden Bundesländern sehr stark variiert.⁸

Zwei weitere Bestandteile niedersächsischer Erwartungshorizonte, die in den anderen beiden Bundesländern nicht vorkommen, sind die explizite Ausweisung der prozentualen Gewichtung der durch die Aufgaben abgedeckten AFB sowie die Benennung von Bewertungskriterien für die Noten »gut« und »ausreichend«.

Damit orientiert sich Niedersachsen sehr stark an den Aufgabenbeispielen, welche mit Erwartungshorizonten in den EPA gegeben werden. Auch dort sind die Kompetenzen zur Lösung der Aufgabenstellungen sowie die unterrichtlichen Voraussetzungen benannt, die Teilaufgaben den AFB zugeordnet und Bewertungskriterien für die Noten »gut« und »ausreichend« definiert.⁹

Ähnlich wie in Bayern und Thüringen wird auch in Niedersachsen der Hinweis gegeben, dass der Erwartungshorizont nur einen möglichen Lösungsweg aufzeigt und »andere Lösungswege mit entsprechender Durchdringungstiefe« denkbar wären. Begründet wird diese Offenheit jedoch durch ein kompetenzorientiertes Bildungsverständnis, indem betont wird, dass die Kompetenzen im Unterricht an unterschiedlichen Inhalten erworben worden sein können. Nicht nur deshalb, sondern auch durch die Benennung der zu prüfenden Kompetenzen weisen die Erwartungshorizonte Niedersachsens am deutlichsten eine Kompetenzorientierung und die Bezüge zu den EPA-Vorgaben aus.

⁷ Dass unter den prozessbezogenen Kompetenzen in allen untersuchten niedersächsischen Aufgabensets »Urteilskompetenz« als eine zu prüfende Kompetenz genannt wird, auch wenn keine explizit auf Urteilsbildung abzielende Teilaufgabe im Set enthalten ist, wurde bereits in Kapitel 6.3.1 dargestellt.

⁸ Eine Abweichung der Aufgabenanzahl in einem Aufgabenset im erhöhten Anforderungsniveau und eine Abweichung in einem Aufgabenset im grundlegenden Anforderungsniveau ergeben sich dadurch, dass in diesen zwei Operatoren in einer Aufgabenformulierungen vorkommen, welche nach der Definition des engen Aufgabenbegriffs als zwei Aufgaben gewertet werden.

⁹ Vgl. EPA, 25–29.

7.2.3 Formale Aspekte thüringischer Erwartungshorizonte

Auch in Thüringen wird den Korrekturhinweisen vorangestellt, dass es sich lediglich um eine Orientierung handelt und dass gleichwertige Lösungsansätze gleichberechtigt sind. Als einziges der untersuchten Bundesländer gibt Thüringen sechs allgemeine, für alle Aufgabenstellungen gültige Bewertungskriterien, wie beispielsweise die »sachliche Richtigkeit« oder »logische Struktur und nachvollziehbare Lösungsstrategie entsprechend der Aufgabenstellung« an. Diese allgemeinen Qualitätsmerkmale sind im ganzen Untersuchungszeitraum gleichgeblieben.

Im Gegensatz zu den aus Bayern und Niedersachsen stammenden Prüfungen sind die Aufgabenstellungen im thüringischen Abitur nicht nummeriert, sondern erscheinen – selbst wenn mehrere Operatoren genutzt werden – als ein Aufgabenkomplex. Dementsprechend sind in den Aufgabenformulierungen für die Prüflinge die in jedem Aufgabenset erreichbaren 60 Bewertungseinheiten zwar benannt, jedoch nicht auf einzelne Teilaufgaben verteilt, wie es in Niedersachsen und Bayern der Fall ist.

Die Erwartungshorizonte beschreiben die Anforderungen der Aufgabe in einigen Fällen zunächst durch einige Sätze und schlüsseln dann die Erwartungen stichpunktartig in einer Tabelle auf, welche nach Operatoren gegliedert ist, die in vielen Fällen nicht deckungsgleich mit den in der Aufgabenstellung benannten Operatoren sind, wie in den Fallanalysen illustriert wird (vgl. Kapitel 7.5).

Darüber hinaus fällt auf, dass in thüringischen Abituraufgabensets anders als in Bayern und Niedersachsen für den AFB I keine eigene Teilaufgabe formuliert wird, welche in bayerischen und niedersächsischen Abiturprüfungen in der Zusammenfassung zentraler Aussagen des dem Aufgabenset zugrunde liegenden Materials konventionalisiert ist.¹⁰ Nur in drei der zwölf aus Thüringen analysierten Aufgabensets ist eine Aufgabenstellung zu finden, die einen zusammenfassenden und reproduzierenden Charakter hat ([TH, 2015, eA (1)], [TH, 2017, eA (1)] und [TH, 2018, eA (2)]). In den anderen Aufgabenstellungen sind die Reproduktionsanforderungen aus AFB I in die komplexen Aufgabenstellungen integriert. So wird beispielsweise eine Reproduktion zur Aufgabe, sich mit der Theodizeefrage auseinanderzusetzen und dabei das Material einzubeziehen, im Erwartungshorizont unter dem Operator »Darstellen/Analysieren« verlangt, ohne dass dies dem Prüfling explizit ersichtlich ist [E-TH, 2016, eA (1)].

¹⁰ Eine solche textzusammenfassende Aufgabenstellung aus AFB I findet sich in allen 24 Aufgabensets der Bundesländer Bayern und Niedersachsen als Einstiegsaufgabe wieder; mit einer einzigen Ausnahme ([BY, 2018 (1)]), in der sich die Reproduktionsaufgabe auf eine Beschreibung eines Gemäldes bezieht.

7.2.4 Zusammenfassung

Ein Vergleich der formalen Aspekte zeigt vor allem Unterschiede in der Transparenz der Anforderungen für die Prüflinge. Dies gilt zum einen in Bezug auf die zu erwartenden Aufgabenstellungen: Die Schülerinnen und Schüler in Niedersachsen wissen durch die starke Formalisierung bereits vorher, wie viele Aufgabenstellungen sie zu erwarten haben und welchem AFB jede Aufgabenstellung zugeordnet sein wird, während dies in Bayern (mit Ausnahme der ersten Aufgabe, die konsequent aus dem AFB I stammt) und in Thüringen nicht deutlich ist. Zum anderen gilt dies in Bezug auf die Wertung der einzelnen Aufgaben: Während den Prüflingen in Thüringen unklar bleibt, wie die 60 erreichbaren Bewertungseinheiten auf die geforderten Aufgaben verteilt werden, wird dies den Prüflingen in Niedersachsen und Bayern vorab bereits transparent gemacht.

Insgesamt betrachtet verfolgen die Erwartungshorizonte von ihrer Anlage her synthetische Beurteilungen, die auf einen Gesamteindruck der in einer Aufgabe geforderten Leistung angelegt sind. Zwar werden einzelne Teilaspekte, die zu beurteilen sind, genannt, es gibt jedoch keine systematische, an Punkte gekoppelte Zusammensetzung dieser, die eine analytische Bewertung ermöglichen würden (vgl. zu synthetischer und analytischer Bewertung Kapitel 3.3.2).

7.3 Fallanalysen aus Bayern

Die folgenden drei Fallbeispiele aus Bayern erschienen bei der Kategorienanalyse – jeweils aus unterschiedlichen Perspektiven – als interessant und in einzelnen Aspekten auch als für die Prüfung von Urteilsfähigkeit im bayerischen Abitur im Fach Evangelische Religionslehre charakteristisch. Sie sollen daher vertieft betrachtet und diskutiert werden, wobei das Besondere und das Charakteristische herausgearbeitet wird.

7.3.1 Die Frage nach dem Maßstab

Im Jahr 2015 wurde als eine von insgesamt sieben Aufgaben im bayerischen schriftlichen Abitur folgende Aufgabenstellung formuliert:¹¹

»2.2 Nehmen Sie zu einer von Grafs zentralen Aussagen kritisch Stellung!« [BY, 2015 (4)]

¹¹ Zum zugrunde liegenden Aufgabenbegriff vgl. Kap. 5.2.1.

Der Aufgabe liegt ein Interview mit dem evangelischen Theologen WILHELM FRIEDRICH GRAF über das Verhältnis von Kirche und Staat zugrunde, welches in der Zeitschrift *zeitzeichen* veröffentlicht wurde und welches in zwei vorangegangenen Teilaufgaben bereits in fünf Thesen zusammengefasst und ins Verhältnis zu Luthers Zwei-Regimente-Lehre und zu Barths Theorie der Königsherrschaft Christi ins Verhältnis gesetzt werden sollte.¹² Die Fragestellung der Urteilsbildung kann vom Prüfling frei gewählt werden, wobei eine Fragestellung zum Verhältnis von Kirche und Staat nahe liegt, da dieses Thema das Interview dominiert. Für die Teilaufgaben können 10 von insgesamt 100 Bewertungseinheiten erreicht werden, das heißt die Lösung macht nur 1/10tel der Bewertung aus. Da die Bearbeitungszeit des Aufgabensets 210 Minuten beträgt, stehen bei einer der Wertung entsprechenden Aufteilung der Zeit dem Prüfling lediglich 21 Minuten zur Bearbeitung der Aufgabe zur Verfügung.

Der Erwartungshorizont gibt ausschließlich folgende Hinweise zur Bewertung:

»Die für die Stellungnahme gewählte Aussage kann mehr oder weniger komplex sein. In der kritischen Stellungnahme kann die gewählte Aussage Grafs abgelehnt oder unterstützt werden; wichtig ist dabei eine schlüssige und differenzierte Begründung.«¹³ [E-BY, 2015 (4)]

Interessanter als die gegebenen Korrekturhinweise ist in diesem Fallbeispiel das, was unangesprochen bleibt. Inhaltlich ist die Aufgabenstellung sehr offen gestellt, so dass kein konkretes Wissen benannt wird, welches zur Lösung der Aufgabe vorausgesetzt wird. Durch den hohen Grad an Offenheit in Bezug auf die Thesenwahl ergeben sich verschiedene Möglichkeiten an Komplexitätsgraden, wie im Erwartungshorizont reflektiert wird (»Die [...] gewählte Aussage kann mehr oder weniger komplex sein«). Offen bleibt jedoch, wie mit den Differenzen hinsichtlich der Komplexität der zu erwartenden Lösung bewertungstechnisch umgegangen werden soll: Wenn die Auswahl der Fragestellung offen sein soll, wie es die Aufgabenstellung formuliert, dann darf die Wahl der Fragestellung nicht von vornherein die zu erreichende Punktzahl determinieren. So muss es

¹² Auch an dieser Stelle zeigt sich nochmals, wie breitgefächert der von den Prüflingen in Bayern verlangte Stoff in einem Abituraufgabenset ist. Es folgen Teilaufgaben, welche die Erläuterung des evangelischen Menschenbildes unter besonderer Berücksichtigung des Freiheitsgedankens, eine Erörterung einer selbstgewählten ethischen Fragestellung aus christlicher Perspektive unter der Berücksichtigung protestantischer Vielfalt und Uneindeutigkeit sowie das Entfalten eines Beispiels »für das Reden von Gott in der modernen Theologie« fordern (vgl. BY, 2015 [4]).

¹³ Kursivsetzung wie im Original.

möglich sein, dass auch die Bearbeitung einer weniger komplexen Fragestellung zu einer Bewertung mit voller Punktzahl führt.

Auch im Hinblick auf die Bewertungskriterien wirken sich die zulässigen Unterschiede im Komplexitätsgrad problematisch aus: Die im Erwartungshorizont benannten Kriterien der Bewertung, die sich dem Bereich *Fähigkeiten und Fertigkeiten* zuordnen lassen, da sie auf den urteilenden Umgang mit dem Wissen zielen, sind »Schlüssigkeit« und »Differenziertheit«. Beide sind maßgeblich von der Komplexität der gewählten Fragestellung abhängig: Eine komplexe Fragestellung bedarf im Vergleich zu einer weniger komplexen Fragestellung einer umfangreicheren Betrachtung, ehe man sie als »differenziert betrachtet« einschätzen würde. Somit ist der Maßstab zur Einschätzung der Ergebnisse an die Auswahl der Fragestellung gekoppelt, was die Einschätzung, ob etwas differenziert ist, beliebig macht.

Im Bereich *Einstellungen und Haltungen zum Gegenstand* wird explizit erwähnt, dass sowohl eine Unterstützung als auch eine Ablehnung der gewählten These möglich ist. Die Offenheit der Aufgabe erschwert eine Konkretisierung möglicher Einstellungen und Haltungen – eine Problematik, die generell mit offenen Aufgaben verbunden ist:

»[J]e offener und komplexer die Aufgabenstellung, desto schwieriger wird die Generierung guter Korrekturanleitungen. Je mehr Entscheidungsspielraum die Prüflinge haben, desto abstrakter werden die Kriterien der Korrekturanleitungen, die dann in vielfältiger Verzweigung beispielhaft ausdifferenziert werden müssen.«¹⁴

In dem vorliegenden Beispiel wurde auf eine beispielhafte Ausdifferenzierung verzichtet. Bewertungskriterien werden nur zwei genannt (schlüssig und differenziert). Diese bleiben – wie oben gezeigt wurde – auf einer abstrakten Ebene. Wie der Zugriff auf die selbstgewählte Fragestellung erfolgen kann – beispielsweise durch Folgenabschätzungen oder Multiperspektivität –, ist nicht weiter bestimmt. Der Erwartungshorizont dieser Aufgabenstellung gibt somit wenig Orientierung für die korrigierende Lehrkraft und genügt nicht, um die für die Vergleichbarkeit der Abschlüsse sehr zentrale Auswertungsobjektivität zu sichern (vgl. Kapitel 3.3.2).

7.3.2 Die Frage nach Bewertungskriterien

Die aus Bayern stammenden Erwartungshorizonte zu Urteilsbildungsaufgaben sind mitunter sehr kurz, wie bereits am vorangegangenen Fall sichtbar wurde

¹⁴ JULIANE KÖSTER, Leistung und Qualität von Korrekturanleitungen im Zentralabitur Deutsch, in: Der Deutschunterricht 1 (2010), 31–38, hier 32.

und wie auch folgendes Fallbeispiel zeigt. Die nun fokussierte Aufgabenstellung fragt nach Dankbarkeit und Demut in Leidsituationen:

»5. Dankbarkeit und Demut (vgl. Material 1, Z. 18–20) auch im Leid? Entfalten Sie dazu eine Antwort und greifen Sie dabei auf einen Ansatz zur Theodizee zurück!« [BY, 2015 (3)]

Auf die Aufgabe entfallen 15 von insgesamt 100 Bewertungseinheiten, wodurch rein rechnerisch 31,5 Minuten der insgesamt 210-minütigen Bearbeitungszeit für diese Aufgabe zur Verfügung stehen. Der Aufgabenstellung liegt kein Text zugrunde, da zwar die Stichworte Dankbarkeit und Demut einem dem Aufgabenset beigelegten Material entstammen, dieses jedoch keinen konkreten Bezug zur Fragestellung nach dem Leid oder der Theodizee-Frage aufwirft. Es handelt sich also um eine der insgesamt nur zwei Aufgabenstellungen des Datenkorpus ohne Materialbezug. Die Fragestellung der Urteilsbildung, ob eine Haltung der Dankbarkeit und Demut auch im Leid angemessen ist, mutet zunächst sehr persönlich und ohne eine konkrete Situation abstrakt an. Durch die geforderte Bezugnahme zur Theodizeefrage wird sie zu einer theologischen Fragestellung, wobei sie persönlich bleibt, da die eigene Sichtweise auf eine Haltung unter Berücksichtigung des Gottesbildes erfragt wird (Welche Auffassung habe ich vom Handeln und Eingreifen Gottes und welche Bedeutung messe ich diesem bei?). Da es sich um eine Frage mit metaphysischem Bezug handelt, die sich auf persönliche Überzeugungen bezieht, treten Kategorien ethisch-philosophischer Urteilsbildung wie gesellschaftliche Werte, rechtliche Perspektiven oder Folgenabschätzungen in den Hintergrund (vgl. Kapitel 2.1). Wie THERESA SCHWARZKOPF herausgearbeitet hat, kommt es in Bezug auf metaphysische Fragen vor allem auf die inhaltliche Kohärenz eines vertretenen Konstrukts und die begründete Abgrenzung gegenüber anderen Konstrukten an (vgl. Kapitel 4.3).

Der Erwartungshorizont trifft folgende Aussagen:

»5 Die Schülerinnen und Schüler sollen die Frage, inwieweit die Haltung von Dankbarkeit und Demut nicht nur im Glück, sondern auch im Leid angemessen ist, auf die Theodizeefrage beziehen. Auf der Basis eines frei gewählten Ansatzes sollen sie eine eigene Antwort formulieren. Möglich ist z. B. der Rückgriff auf biblische Traditionen wie Hiob oder die Passionsüberlieferung; geeignet ist auch ein philosophischer (z. B. Leibniz) oder theologischer Ansatz (z. B. Luther).«¹⁵ [E-BY, 2015 (3)]

Auch dieser Erwartungshorizont lässt die Bewertungskriterien weitestgehend offen. Die Fragestellung wird durch den Erwartungshorizont paraphrasiert, oh-

¹⁵ Formatierung wie im Original.

ne dass dabei zentrale Verstehens- und Bearbeitungsschritte benannt würden, welche zur Lösung der Aufgabe notwendig sind. Der Erwartungshorizont setzt die Verbindung zwischen der Theodizeefrage und einer Haltung von Dankbarkeit und Demut in Leidsituationen als zentral. Ein inhaltlicher Zusammenhang zwischen beiden Bereichen wird jedoch vom Erwartungshorizont selbst nicht hergestellt. So bleibt im Bereich *Wissen* offen, welche Aspekte der Theodizeefrage als Verstehensvoraussetzung gelten müssen und wie ein Brückenschlag zu den angesprochenen Traditionen und Konzeptionen aussehen kann.

Ebenso unkonkret bleibt der Erwartungshorizont im Bereich *Fähigkeiten und Fertigkeiten* und *Einstellungen und Haltungen* zum Gegenstand, da er keine Qualitätsmerkmale benennt, mit denen die erwartete Antwort auf die Fragestellung zu bewerten wären oder wie mögliche persönliche Sinnkonstruktionen zu diagnostizieren sind.

Die geringe Zahl an Bewertungseinheiten, die daraus abgeleitete knappe Zeitvorgabe und der ebenso knappe Erwartungshorizont legen die Annahme nahe, dass die Aufgabenstellung in ihrer Komplexität in der Erstellung des Aufgabensets unterschätzt wurde.

7.3.3 Die Frage nach der Ausformulierung von Pro- und Kontrapositionen

Wie durch die Analysekategorie H *Formulierung konträrer Positionen in den Erwartungshorizonten* gezeigt werden konnte, bleiben die möglichen Positionen in Hinblick auf die Fragestellung häufig unbenannt. Dadurch kommt die Kontroversität der zu beurteilenden Fragestellung in den Erwartungshorizonten nicht in den Blick, wie exemplarisch an einer Aufgabenstellung aus Bayern aus dem Jahr 2019 gezeigt werden kann. Dort wurde die Aufgabe gestellt:

»5 Setzen Sie sich damit auseinander, inwieweit grundlegende Aussagen des Apostolischen Glaubensbekenntnisses »die letzten Fragen nach dem Sinn unserer Existenz, dem Warum, Woher, Wozu« (Z. 40 f.) umfassend beantworten können.« [BY, 2019 (4)]

Der dem Aufgabenset zugrunde liegende Text stellt die Frage nach der Religion in der säkularen Moderne im Kontext staatlichen Handelns in den Mittelpunkt. Die in der Aufgabe aufgeworfene Fragestellung wird in der Textgrundlage nicht behandelt. Die Kenntnis des Apostolischen Glaubensbekenntnisses wird bei den Schülerinnen und Schülern vorausgesetzt. Es ist nicht als Materialgrundlage

gegeben.¹⁶ Auf die Aufgabe entfallen 20 von 100 Bewertungseinheiten des Aufgabensets, welches insgesamt neun Aufgaben enthält.¹⁷ 20 % der Bearbeitungszeit entsprechen 42 Minuten, welche den Prüflingen rechnerisch zur Bearbeitung der Aufgabe zur Verfügung stehen. Beurteilt werden soll die Tragweite des Apostolischen Glaubensbekenntnisses in Bezug auf die Beantwortung »letzter Fragen«. Der vollständige Erwartungshorizont der Aufgabe sieht folgendermaßen aus:

»Eine Auseinandersetzung muss zunächst grundlegende Aussagen des Apostolikums benennen und sie als Antworten auf die ›letzten Fragen‹ des Textes kenntlich machen:

- Gott der Vater als Schöpfer und Erhalter der Welt;
- Jesus Christus als Sohn Gottes wird Mensch, um das Reich Gottes bereits zu seinen Lebzeiten und durch sein Wirken aufzurichten, indem er Heil stiftet;
- Jesus Christus, der den Tod als Erster überwunden hat und die Hoffnung auf dessen endgültige Überwindung und das ewige Leben begründet;
- Jesus Christus, der bei seiner Wiederkunft die Welt endgültig heilen wird;
- der Heilige Geist, der Trost spendet und Menschen untereinander und zur Gemeinschaft in und mit Gott verbindet.

Fragen wie etwa die nach dem Leid, dem Gefühl der Abwesenheit Gottes, des Hingeworfenseins und der Fragmentarität stehen hingegen weniger unmittelbar im Fokus des Apostolikums.

*Es ist eine strukturierte Darstellung zu erwarten, die in ein begründetes Urteil je nach Schwerpunktsetzung mündet.*¹⁸ [E-BY, 2019 (4)]

Der Erwartungshorizont fokussiert im Bereich *Wissen* ausschließlich die grundlegenden Aussagen des Apostolikums, ohne diese explizit auf die Fragestellung zu beziehen. Die Aussagen des Apostolikums zu kennen, ist eine Voraussetzung zur Lösung der Aufgabenstellung, die sich als eine Reproduktion von Wissen dem AFB I zuordnen lässt. Die eigentliche Fragestellung nach einer Stellungnahme zur Tragfähigkeit des Apostolikums in Bezug auf die letzten Fragen gehört zu AFB III. Die zentralen Fragestellungen, die der Aufgabe ihre Kontroversität verleihen, wie beispielsweise was eine »umfassende Beantwortung letzter Fragen« ausmachen würde und welche Prämissen man akzeptieren muss, damit das Apostolikum in Bezug auf letzte Fragen Antworten geben kann, werden im Erwartungshorizont

¹⁶ Im Untersuchungszeitraum wurde in Thüringen ebenso eine Urteilsbildungsaufgabe gestellt, in der das Apostolikum eine Rolle spielt [TH, 2018, eA (1)]. In dieser wurde das Glaubensbekenntnis als Materialgrundlage gegeben. In keinem der beiden Bundesländer wird das Apostolikum im Lehrplan erwähnt.

¹⁷ Definiert nach einem engen Aufgabenbegriff (vgl. Kap. 5.2.1).

Die Verteilung der Bewertungseinheiten ist den Prüflingen ersichtlich.

¹⁸ Hervorhebungen wie im Original.

nicht reflektiert. Unterschiedliche Positionen in Bezug auf die Fragestellung werden nicht genannt. Deutungsspielraum lässt die Formulierung, dass verschiedene Aspekte »hingegen weniger unmittelbar im Fokus des Apostolikums« stehen, da zwei Lesarten des Satzes möglich sind: Zum einen kann er andeuten, dass ein Eingehen auf diese Aspekte als falsch bewertet werden soll, da sie sich zu weit vom Apostolikum entfernen. Zum anderen kann er als ein Argument gegen die These gelesen werden, das Apostolikum könne die grundlegenden »letzten Fragen« umfassend beantworten.

Im Bereich *Fähigkeiten und Fertigkeiten* sind auch in diesem Erwartungshorizont nur zwei Qualitätsmerkmale benannt (strukturierte Darstellung und Vorhandensein einer Begründung), anhand derer die Urteilsbildung bewertet werden kann. Obwohl es sich auch hier wieder um eine Fragestellung handelt, die grundlegend die persönlichen Wirklichkeitsannahmen und das persönliche Glaubensbekenntnis betreffen, sind im Bereich *Einstellungen und Haltungen* zum Gegenstand der Urteilsbildung im Erwartungshorizont keine Hinweise enthalten. Die Herausforderungen der geforderten Urteilsbildung werden in diesem Erwartungshorizont nicht explizit.

7.4 Fallanalysen aus Niedersachsen

7.4.1 Ausformulierung von Pro- und Kontrapositionen

Dass die Ausformulierung verschiedener Positionen in der für Erwartungshorizonte gebotenen Kürze möglich ist, kann an einem Erwartungshorizont aus Niedersachsen zur Aufgabenstellung

»3. Entwerfen Sie einen Brief an Nikolaus Schneider zu dessen Sicht des Menschen.«
[NI, 2016, gA (1)]

gezeigt werden. Das Aufgabenset umfasst drei Aufgabenstellungen. Die Gewichtung der Aufgabe ist mit 30 % angegeben. Bemisst man die Zeit adäquat zur Gewichtung, stehen dem Prüfling 66 Minuten – also etwas mehr als eine Stunde – der insgesamt 220-minütigen Bearbeitungszeit zur Lösung dieser Aufgabenstellung zur Verfügung. Die von NIKOLAUS SCHNEIDER, evangelischer Theologe und ehemaliger Ratsvorsitzender der EKD, im Text vertretene Anthropologie definiert sich maßgeblich durch den Anspruch Gottes an den Menschen. Die von ihm benannten zentralen Aspekte des Menschseins sind daher auf die menschliche Abhängigkeit von Gott bezogen. Die Fragestellung der Urteilsbildung ist durch die christliche Sicht SCHNEIDERS eine theologische, da sie nach Ablehnung oder Zustimmung einer (von mehreren möglichen theologischen) Sichtweise(n) auf den Menschen fragt.

Da sich die beiden vorangehenden Aufgaben dieses Aufgabensets bereits mit dem Text SCHNEIDERS befassen, indem zuerst der Inhalt wiedergegeben und dann die Aussagen SCHNEIDERS anhand biblisch-theologischer Grundlagen christlicher Ethik entfaltet werden sollen, werden die Prüflinge in Bezug auf die Urteilsbildungsaufgabe bereits »vorentlastet«. »Vorentlastet« ist dadurch auch der Erwartungshorizont der Aufgabe: Dass er auf der Ebene der Aufzählung von Grundaussagen verhaftet bleibt, wie es im vorangegangenen Beispiel der Fall war (vgl. Kap. 7.2.3), ist hier nicht möglich, da diese Ebene bereits in den vorangegangenen Aufgaben Betrachtung fand. Im Zentrum der Erwartungsformulierungen zu oben zitierter Aufgabe steht die Urteilsbildung. Dort heißt es in vollständiger Wiedergabe:

»Diese Teilaufgabe ist schwerpunktmäßig dem AFB III zuzuordnen.

Zum einen erfordert die Aufgabenstellung die Beachtung des formalen Rahmens eines Briefes; zum anderen sollen sich die Prüflinge kreativ mit der auf den Text bezogenen Aufgabenstellung auseinandersetzen.

Die Aufgabe überlässt den Prüflingen u. a. die Entscheidung,

- sich gegen die Sicht Schneiders zu positionieren und z. B. einen anderen anthropologischen Entwurf darzustellen (z. B. Verantwortung und Freiheit des Menschen auch ohne Gottesglauben),
- andere biblisch-christliche Akzente des Menschseins herauszustellen bzw. anders zu gewichten, als Schneider dies tut, z. B. Betonung der Gnade Gottes gegenüber dem Menschen oder die Rechtfertigung durch den Glauben,
- einzelne Aspekte des Textes besonders in den Fokus zu nehmen, wie z. B. den, dass dem Menschen das für ein gelingendes Leben Entscheidende von Gott gesagt werde, den der Gefahr der Selbstüberschätzung oder den der häufig schwierigen Entscheidungsfindung angesichts der Probleme in der Welt, [sic!]

Im Brief sollten der Bezug zum Text Schneiders und eine eigene Position des Prüflings deutlich werden.« [E-NI, 2016, gA (1)]

Unter der Überschrift »2.2. Bewertungskriterien für die Note »gut« bzw. »ausreichend« ist darüber hinaus zur Aufgabe formuliert:

»Eine Leistung kann mit »gut« bewertet werden, wenn

- [...]
- der Brief überzeugend und einfallsreich gestaltet ist und eine ausführliche, aspektreiche und nachvollziehbare Bezugnahme zu Schneiders Menschenbild darstellt. [...]

Eine Leistung kann mit »ausreichend« bewertet werden, wenn

- [...]
- die Ausführungen im Brief zu Schneiders Sicht des Menschen insgesamt nachvollziehbar dargestellt sind. [...]« [E-NI, 2016, gA (1)].

Der Erwartungshorizont benennt im Bereich *Einstellungen und Haltungen zum Gegenstand der Urteilsbildung* verschiedene Positionen und gibt in knappen Stichpunkten Beispiele, wie diese Positionen inhaltlich gefüllt werden könnten. Betrachtet man die Voraussetzungen der Aufgabe im Bereich *Wissen* sind eine Ablehnung der Aussage und eine begrenzte Zustimmung voraussetzungsreicher als eine Zustimmung, da die Kenntnis oder Erarbeitung einer anderen in sich kohärenten Anthropologie notwendig ist, während man bei einer Zustimmung in erster Linie weitere Argumente für das dargestellte Konstrukt anführen muss.

Als zur Lösung der Aufgabenstellung erforderlichen *Fähigkeiten und Fertigkeiten* werden die formale Gestaltung eines Briefes, die eigene Positionierung sowie die Bezugnahme zum Text SCHNEIDERS genannt. In Hinblick auf Qualitätsmerkmale lassen sich dem Erwartungshorizont folgende Stichworte entnehmen:

Umgang mit Aufgabenstellung:

- kreativ

Gestaltung des Briefes:

- überzeugend
- einfallsreich[]

Bezugnahme zu SCHNEIDERS Menschenbild/Urteilsbildung:

- ausführlich
- aspektreich
- nachvollziehbar
- eigene Positionierung mit eigener Schwerpunktsetzung.

Im Gegensatz zu vielen anderen Erwartungshorizonten aus Niedersachsen, die den Umfang der Argumente explizit nivellieren, wird in diesem Erwartungshorizont die Ausführlichkeit der Bezugnahme zu SCHNEIDERS Menschenbild als ein Qualitätskriterium explizit genannt. Die benannten Qualitätskriterien legen einen binären Beurteilungsmodus zwischen »vorhanden/erfüllt« oder »nicht vorhanden/nicht erfüllt« nahe, wobei sie auch diesbezüglich einen hohen Deutungsspielraum zulassen (Was muss gegeben sein, damit der Umgang mit der Aufgabe als »kreativ« bezeichnet werden kann und ab wann wird aus »kreativ« das Urteil »Aufgabenstellung verfehlt«? Woran misst sich Ausführlichkeit?). Eine Graduierung der Merkmalsausprägung, wie sie die Vergabe von Bewertungseinheiten erfordert, ist der korrigierenden Lehrkraft überlassen, so dass der Erwartungshorizont keine Auswertungsobjektivität sicherstellt.

Dennoch sticht der Erwartungshorizont durch die Benennung verschiedener Positionen und Qualitätsmerkmale im Untersuchungskorpus positiv hervor, da er die Herausforderung der Urteilsbildung fokussiert und auf die Fragestellung hin ausformuliert.

7.4.3 Urteilsbildung zu gegenwartsbezogenen gesellschaftspolitischen Fragestellungen

Dass die Abiturprüfungsaufgaben gesellschaftspolitische Fragestellungen aufnehmen, lässt sich in Zusammenhang mit den weltweit erhöhten Migrationsbewegungen im Jahr 2015 zeigen: Alle drei untersuchten Bundesländer greifen das Thema Migration im Abitur auf: Im bayerischen Abitur wurde 2016 ein Aufgabenset gestellt, dem eine Ansprache des evangelischen Theologen HEINRICH BEDFORD-STROHMS über die verstärkte Migration zugrunde liegt.¹⁹ Eine darin enthaltene Aufgabenstellung fordert die Beurteilung der Veröffentlichung eines im September 2015 entstandenen Fotos aus Sicht einer christlichen Medienethik [BY, 2016 (3)]. Das Foto hatte Kontroversen über die Verantwortung der Presse ausgelöst, da ein syrisches, auf der Flucht gestorbenes und an den türkischen Stand gespültes Kind darauf zu sehen war.

Im thüringischen Abitur wird die Thematik ebenfalls 2016 aufgegriffen. Dort sollen die Prüflinge auf Grundlage verschiedener Materialien die biblische Botschaft in Bezug auf das Fremde und auf Flüchtlinge herausarbeiten sowie unter Einbezug des christlichen Menschenbildes einen »offenen Willkommensbrief an Flüchtlinge [entwerfen], die in unser Land kommen« [TH, 2016, eA (2)]. Eine Aufgabe zur Urteilsbildung ist in diesem Aufgabenset nicht enthalten.

Im niedersächsischen Abitur wird bereits 2015 in einer Aufgabenstellung auf die Thematik Flucht Bezug genommen [NI, 2015, eA (2)]. Das Aufgabenset steht unter der Überschrift »Ist Glauben Privatsache?«. Als Texte liegen dem Aufgabenset ein im Magazin *zeitzeichen* erschienener Artikel der evangelischen Theologin PETRA BAHR zum Öffentlichkeitsanspruch der Kirchen zugrunde sowie ein Aufruf des Vorstands der Arbeitsgemeinschaft christlicher Kirchen (ACK) Bayern zum konkreten Einsatz für Flüchtlinge. In einer der insgesamt vier Aufgabenstellungen wird eine Urteilsbildung von den Prüflingen gefordert. Dort heißt es:

»4. Nehmen Sie in einem Brief an den Vorstand der ACK Bayern einen begründeten Standpunkt zu deren Aufruf ein (M2). Beziehen Sie in Ihre Ausführungen ausgewählte Aspekte aus dem Text von Petra Bahr ein.« [NI, 2015, eA (2)]

Die Gewichtung der Aufgabe ist mit 30% angegeben, so dass rechnerisch 90 Minuten Bearbeitungszeit der insgesamt 300-minütigen Gesamtbearbeitungszeit auf diese Aufgabe entfallen. Der Einbezug des Standpunktes von BAHR zum Verhältnis von Kirche und Öffentlichkeit ist bereits durch die drei vorangestellten Aufgaben vorstrukturiert. Dies trifft auf die Auseinandersetzung mit

¹⁹ HEINRICH BEDFORD-STROHM ist seit 2014 – also auch zum Zeitpunkt der zitierten Rede – Ratsvorsitzender der EKD.

der Position des ACK Bayern nicht zu – die Thematik Flucht rückt erst durch diese (letzte) Aufgabe in den Fokus. Der Erwartungshorizont gibt folgende Bewertungshinweise zur Aufgabenstellung:

»Diese Aufgabe ist schwerpunktmäßig dem AFB III zuzuordnen.

Die ACK Bayern ruft zu einem differenzierten Engagement für Flüchtlinge auf: Ausgehend von dem in der Bibel geforderten zugewandten Umgang mit den Fremden, der Grunderfahrung der Befreiung aus Ägypten sowie der Predigt Jesu über die Nächstenliebe sei ein engagiertes Handeln der Christen angesichts der Flüchtlingsproblematik erforderlich. Dieses beinhaltet konkret soziale Beziehungen, materielle Unterstützung, das Aufgeben von Vorurteilen sowie die Arbeit an politischen Strukturen, so dass ein geschwisterlicher Umgang und eine Wahrnehmung der Verantwortung möglich würden.

Die Prüflinge müssen sich in einem Brief zu diesen Forderungen verhalten und dabei begründen, ob sie ihnen zustimmen, sie als zu weitgehend oder als zu schwach bewerten würden. Dabei sind Bezüge zu den Ausführungen von Petra Bahr herzustellen, wie z.B. der Einsatz der Kirche für die Schwachen, ihr Wächteramt, die Evangeliumsgemäßheit dieser Verantwortung, das christliche Menschenbild oder die Verantwortung aller Christen. Auch könnte die von Bahr geforderte, in dem Aufruf aber nicht beachtete Verbindung von privatem und öffentlichem Glauben mit in die Auseinandersetzung einbezogen werden.

Die Form des Briefes sollte u. a. durch den Adressatenbezug, das Eingehen auf die Inhalte des Aufrufs und das Herausarbeiten des eigenen Standpunktes deutlich werden. Auch können besondere Stilmittel wie z. B. rhetorische Fragen verwendet werden.

[...]

2.2. Bewertungskriterien für die Noten »gut« und »ausreichend«

Eine Leistung kann mit »gut« bewertet werden, wenn [...]

Der Brief an den Vorstand entwickelt, ausgehend von dessen Forderungen und mit sachgemäßem Rückgriff auf die Ausführungen von Petra Bahr, einen eigenen, begründeten Standpunkt. Dabei ist die Form des Briefes in überzeugender Weise berücksichtigt. [...]

Eine Leistung kann mit »ausreichend« bewertet werden, wenn [...]

Der Brief an den Vorstand bezieht sich in Teilen auf dessen Forderungen und auf die Ausführungen von Petra Bahr. Er lässt in Ansätzen einen eigenen Standpunkt sowie die Form eines Briefes erkennen. [...]« [E-NI, 2015, eA (2)]

In Bezug auf das zur Lösung der Aufgabe notwendige *Wissen* klärt der Erwartungshorizont zunächst die Position des ACK Bayern, die die Prüflinge herausarbeiten müssen, um sie beurteilen zu können. Die geforderte Bezugnahme auf den Text BAHRS ist mit inhaltlichen Beispielen unterlegt.

In Bezug auf den Bereich *Einstellungen und Haltungen* werden drei verschiedene mögliche Stellungnahmen benannt, die jedoch nicht mit einzelnen Argumenten unterlegt sind. In der Trias »Zustimmung«, »zu weitgehend«, »oder

zu schwach« wird eine grundsätzliche Ablehnung der zu beurteilenden Position nicht in Erwägung gezogen.

Fähigkeiten und Fertigkeiten beziehen sich auf die Umsetzung der Briefform, welche durch »den Adressatenbezug, das Eingehen auf die Inhalte des Aufrufs und das Herausarbeiten des eigenen Standpunktes« konkretisiert werden. Die Verwendung stilistischer Mittel wird als optional gekennzeichnet. In Bezug auf Urteilsbildung sind die Angaben eher rar: Bis auf einen »sachgemäße[n] Rückgriff auf die Ausführungen von Petra Bahr« werden keine Qualitätsmerkmale genannt.

7.4.4 Exkurs: Fehlende Schwierigkeitsdifferenzierung der Anforderungsniveaus im Bereich Urteilsfähigkeit

Vergleicht man die aus Niedersachsen stammenden Urteilsbildungsaufgaben auf erhöhtem Anforderungsniveau mit denen auf grundlegendem Anforderungsniveau wird deutlich, dass in Bezug auf Urteilsbildungsaufgaben keine Schwierigkeitsdifferenzierungen vorgenommen werden. Dies wird besonders in den Jahren 2017 und 2019 offensichtlich, in denen die gleiche Aufgabenstellung zur Urteilsfähigkeit sowohl in einem Aufgabenstet mit grundlegendem also auch in einem Aufgabenstet mit erhöhtem Anforderungsniveau gestellt wurde. In beiden Fällen liegt den Aufgabensets der gleiche Text zugrunde. Es handelt sich um folgende Aufgabenstellungen:

3./4. »Setzen Sie sich mit der Position Steffenskys auseinander.« [NI, 2017, gA (1)] / [NI, 2017, eA (2)]

3./4. »Entwerfen Sie einen Kommentar zu Käßmanns friedensethischer Position.« [NI, 2019, gA (1)] / [NI, 2019, eA (1)]

Die Erwartungshorizonte dieser Teilaufgaben, die dem AFB III zugeordnet werden, sind in beiden Fällen jeweils nahezu identisch. Die Anforderungsunterschiede werden im AFB II erzeugt: Während in der Teilaufgabe zu AFB I in beiden Aufgabensets jeweils ebenso die gleiche Aufgabenstellung zur Wiedergabe des Textinhaltes und ein identischer Erwartungshorizont gegeben sind, weichen die Aufgabenstellungen zu AFB II voneinander ab. Zum einen wird im erhöhten Anforderungsniveau eine Teilaufgabe mehr gefordert, wofür den Prüflingen auch eine um 80 Minuten erhöhte Bearbeitungszeit zugestanden wird als in Prüfungen mit grundlegendem Anforderungsniveau.²⁰ Zum anderen werden 2017 auf

²⁰ Die Bearbeitungszeit für eine Prüfung mit grundlegendem Anforderungsniveau beträgt 220 Minuten, während die Bearbeitungszeit für eine Prüfung mit erhöhtem Anforderungsniveau 300 Minuten beträgt.

grundlegendem Anforderungsniveau in Aufgabe zwei lediglich Begriffserläuterungen zu »Schuld« und »Vergebung« gefordert, welche ihren Ausgangspunkt im zugrunde liegenden Text finden sollen. Dagegen sollen auf erhöhtem Anforderungsniveau ebenso anknüpfend an den Text Grundzüge des biblisch-christlichen Menschenbildes entfaltet und in der weiteren Aufgabe die Ausführungen STEFFENSKYS zu »Schuld« und »Vergebung« mit der lutherischen Rechtfertigungslehre verglichen werden. Auch im Fallbeispiel aus dem Jahr 2019 unterscheiden sich die Aufgabenstellungen im AFB II: Im grundlegenden Anforderungsniveau soll ausgehend vom Text die Botschaft Jesu vom Reich Gottes erläutert werden, während im erhöhten Anforderungsniveau Möglichkeiten der Kirche zur gesellschaftlichen Mitverantwortung entfaltet und eine Kunstinstallation ins Verhältnis zu den Aussagen im Text gesetzt werden sollen.

Innerhalb des AFB II ist folglich in beiden Fällen nicht nur eine quantitative Steigerung, sondern auch ein Komplexitätszuwachs der Aufgabenstellung im Vergleich zum grundlegenden Anforderungsniveau zu erkennen, welcher in der geforderten Verknüpfung der Aussagen des Textes mit anderen Wissensbeständen (biblisch-christliches Menschenbild und lutherische Rechtfertigungslehre) beziehungsweise mit der Deutung der Kunstinstallation zum Ausdruck kommt. In Bezug auf Urteilsbildung wird jedoch sowohl im grundlegendem als auch im erhöhten Anforderungsniveau der gleiche Komplexitätsgrad gefordert. Die Erwartungshorizonte zur Teilaufgabe sind 2017 bis auf einen Formulierungsunterschied identisch.²¹ Gleiches gilt für die Erwartungshorizonte des »Doppelgängers« 2019, bei dem im erhöhten Anforderungsniveau lediglich ein Argument auf Pro- und ein Argument auf Kontraseite hinzutreten, welche sich konkret auf die Rolle der Kirche in Hinblick auf eine pazifistische Politik beziehen und damit Impulse aus Aufgabenstellung zwei dieses Aufgabensets aufnehmen.²² Da es sich bei den in den Erwartungshorizonten angegebenen Argumenten jedoch um Beispiele handelt, die keinen verbindlichen Charakter haben, fallen die beiden zusätzlichen Argumente nicht ins Gewicht.

An den beiden »Doppelgängern« kann also gezeigt werden, dass Niedersachsen im Kompetenzbereich Urteilsfähigkeit hinsichtlich der Bewertungskri-

²¹ Es gibt einen Unterschied in der Benennung der Bewertungskriterien für eine »gute« und eine »ausreichende« Benotung, der jedoch lediglich auf Formulierungsebene liegt: Im erhöhten Anforderungsniveau wird für eine gute Leistung gefordert, dass »bei der Stellungnahme fachgemäße Kenntnisse zum Gottesbild analytisch auf die Position Steffenskys bezogen sind und ein eigenständiges Urteil aspektreich begründet ist« [E-NI 2017, eA (2)], während es auf grundlegendem Anforderungsniveau heißt, dass »die Auseinandersetzung mit den Aussagen über Gott fachgerecht und analytisch auf den Text bezogen [...] und das eigene Urteil zum Gottesbild differenziert begründet« sein sollen [E-NI 2017, gA (1)].

²² An diesem Beispiel wird deutlich, dass die Aufgabenstellungen in den Aufgabensets in Niedersachsen aufeinander aufbauend konzipiert sind.

terien keine Differenzierung zwischen grundlegendem und erhöhtem Anforderungsniveau vornimmt, sondern dass alle Prüflinge Urteilsfähigkeit auf dem gleichen Niveau nachweisen müssen. Dies könnte darin begründet sein, dass das Urteilen über religiöse und ethische Fragestellungen als ein Mindeststandard angesehen wird, der auch auf grundlegendem Anforderungsniveau vollumfänglich erreicht werden soll. Gegen diese Erklärung spricht jedoch, dass Urteilsbildung im niedersächsischen Lehrplan für die Oberstufe keine hervorgehobene Rolle einnimmt.²³ Blickt man auf die im gesamten Untersuchungskorpus erhobenen sehr allgemein formulierten Qualitätsmerkmale, liegt der Schluss nahe, dass kein differenziertes Konzept von Urteilsbildung vorliegt, welches die Benennung unterschiedlicher Niveaustufen ermöglichen würde. Zu konstatieren ist, dass der in den EPA formulierte Anspruch, die Unterscheidung in ein grundlegendes und ein erhöhtes Anforderungsniveau soll eine Differenzierung »im Hinblick auf die Komplexität des Stoffes, im Grad der Differenzierung und Abstraktion der Inhalte und Begriffe, im Anspruch an Methodenbeherrschung und in der Selbstständigkeit der Lösung von Problemen«²⁴ vornehmen, in Bezug auf Urteilsfähigkeit nicht eingeholt wird.²⁵

Von der Studie unerfasst bleiben selbstverständlich individuelle Maßstäbe und Erwartungen der Korrektoren, welche sich zwischen den Anforderungsniveaus unterscheiden könnten. Auch hier zeigt sich erneut die Grundproblematik, dass sich die in den Erwartungshorizonten genannten Bewertungskriterien schwer objektivieren lassen. So wird beispielsweise 2019 in beiden Erwartungshorizonten die begründete Darlegung der eigenen Position sowie die Durchdringungstiefe der Argumentation genannt. Diese Formulierung macht es möglich, dass der Korrektor für die Durchdringungstiefe im grundlegenden Anforderungsniveau subjektiv einen anderen Maßstab ansetzt, als wenn er die Durchdringungstiefe im erhöhten Anforderungsniveau bewertet. Damit ist erneut gezeigt, dass die mangelnde Standardisierung der in den Erwartungshorizonten genannten Begriffe eine negative Auswirkung auf die Auswertungsobjektivität hat.

²³ Vgl. NIEDERSÄCHSISCHES KULTURMINISTERIUM (Hrsg.), Kerncurriculum für das Gymnasium. Der Lehrplan orientiert sich ebenso wie die Prüfungsaufgaben und die Erwartungshorizonte stark an den EPA sowie den in den EPA genannten Kompetenzen (vgl. Kapitel 3.2).

²⁴ Vgl. EPA, 11.

²⁵ Wenn dies hier in Bezug auf Niedersachsen formuliert wird, darf nicht unerwähnt bleiben, dass Bayern und Thüringen die beiden Niveaudifferenzierungen in schriftlichen Abiturprüfungen im Fach Evangelische Religionslehre nicht anbieten.

7.5 Fallanalysen aus Thüringen

7.5.1 Komplexität und Offenheit der Aufgabenstellung

Dass die Aufgabensets aus Thüringen nicht durch Nummerierungen gegliedert sind, wie es in Niedersachsen und Bayern der Fall ist, wurde bereits in der Betrachtung formaler Aspekte erwähnt (vgl. Kapitel 7.2.3). Diese fehlende Gliederung kann zu einer erhöhten Komplexität führen, wie man an folgendem Beispiel aus dem Jahr 2014 zeigen kann:

»Entwerfen Sie auf der Grundlage der Materialien einen Essay, in dem Sie sich mit der Botschaft Jesu auseinandersetzen.

Zeigen Sie auf, welche Rolle christliche Maßstäbe für Ihre Entscheidungen spielen.«²⁶
[TH, 2014 (2)]

Die beiden Handlungsaufforderungen der Aufgabenstellung werden nicht miteinander ins Verhältnis gesetzt, so dass unklar bleibt, ob Teilaufgabe zwei in das Essay einfließen soll oder einer separaten Bearbeitung bedarf. Das Aufgabenset umfasst ausschließlich diese beiden Aufgaben, so dass die gesamte 270-minütige Bearbeitungszeit für diese beiden Aufgaben zur Verfügung steht. Eine Verteilung der insgesamt 60 zu erreichenden Bewertungseinheiten auf die beiden Aufgaben wird dem Prüfling nicht ersichtlich.

Die Aufforderung der Auseinandersetzung mit der Botschaft Jesu ist sehr allgemein formuliert. Die Schwerpunkte und Perspektiven für die Auseinandersetzung werden von der Aufgabenstellung nicht vorgegeben, wodurch viele Bearbeitungswege möglich werden. Eine inhaltliche Eingrenzung erfährt die Aufgabenstellung allein durch den geforderten Bezug auf die Materialien. Diese setzen sich aus der Angabe zweier Bibeltexte (der Bergpredigt in Mt 5–7 und das Doppelgebot der Liebe in Mk 12,28–34) (M 1), sowie aus zwei Liedtexten zusammen, wobei der erste Liedtext (M 2) *Jesu, geh voran* von NIKOLAUS LUDWIG VON ZINZENDORF aus dem 18. Jahrhundert und der zweite Liedtext (M 3) des Rocksongs *Wünsch dir was* von den TOTEN HOSEN aus den 1990er Jahren stammt. Der Liedtext der TOTEN HOSEN wird weder durch die Nennung der Band noch durch die Zeit kontextualisiert, während für das Lied *Jesu, geh voran* »Nikolaus Ludwig von Zinsendorf [sic!]« als Verfasser genannt wird. Die Materialien wirken eklektisch zusammengestellt, da sie aus unterschiedlichen Zeiten und Kontexten stammen. Offenbar ist es die Aufgabe der Prüflinge, die Materialien in eine Beziehung zueinander zu setzen, wobei anhand der Aufgabenstellung nicht deutlich wird, ob sie ein Auswahlangebot darstellen oder ob alle Materialien Berücksichtigung finden sollen. Im Erwartungshorizont wird deutlich, dass die Analyse aller Ma-

²⁶ Zeilenumbruch wie im Original.

terialien erwartet wird. Der vollständige Erwartungshorizont sieht folgendermaßen aus:²⁷

»Christologie/Glauben heute – Lehren von Gott

Beim Entwerfen des Essays benennen die Prüfungsteilnehmer inhaltliche Aussagen der Bergpredigt und erschließen diese danach. Sie formulieren den Essay auf der Grundlage der analysierten Materialien und unter Einbeziehung geeigneter biblischer Texte sowie persönlicher Erfahrungen situationsgerecht und adressatenbezogen.

	AB	BE
Benennen/Analysieren:	I/II	10
Materialien 1 und 2 aufgabenbezogen erschließen und in situationsgerechte Zusammenhänge einordnen		
Material 1		
– inhaltliche Aussagen der Bergpredigt benennen		
– Gottesliebe und Nächstenliebe als höchste ethisch-moralische Grundsätze für das Zusammenleben deuten		
Material 2		
– das Leben als Leben in der Nachfolge Jesu deuten		
– Glaubenszusage in schwierigen Situationen aufzeigen		
Interpretieren:	III	10
Liedtext sachgemäß analysieren und auf der Basis der methodisch reflektierten Deutung schlüssig auslegen		
Material 3		
– Hoffnungsgedanken aufzeigen		
– Gedanken eines friedfertigen Miteinanders analysieren		
– eschatologische Vorstellungen herausarbeiten		
Belegen:	II	10
Aussagen mittels geeigneter Textstellen aus der Bibel stützen und in den Essay einbeziehen		
Entwerfen und Gestalten:	III	30
einen Essay kreativ sowie sach- und situationsgerecht entsprechend der Aufgabe formulieren		

²⁷ Die Formatierung wurde weitestgehend beibehalten. Einige Zeilenabstände und die Schriftgröße wurden verringert.

-
- Essay auf Grundlage der analysierten Materialien unter Einbeziehung geeigneter biblischer Texte sowie persönlicher Erfahrungen situationsgerecht und adressatenbezogen entwerfen
 - in kreativer und offener Schreibform eine begründete Haltung zum Thema darstellen
 - in pointierter Form Denkanstöße schaffen
-

[E-TH, 2014 (2)]

Das Aufgabenset beinhaltet zwei Handlungsanweisungen, die durch die Operatoren »entwerfen« und »aufzeigen« benannt sind. Wie bereits im Kapitel »Grenz- und Sonderfälle« dargestellt wurde, entspricht die Definition des Operators »zeigen Sie auf« (AFB I) nicht der inhaltlich formulierten Anforderung, die in AFB III liegt (vgl. Kapitel 5.2.3). Während dies vor allem eine Fehlpassung mit den AFB der EPA darstellt, die die Prüflinge eventuell gar nicht bemerken, führt die fehlende Passung zu den Operatoren des Erwartungshorizonts zu intransparenten Schwerpunktsetzungen: Im Erwartungshorizont der Aufgabe wird der Operator »aufzeigen« nicht aufgegriffen. Stattdessen ist der Erwartungshorizont durch die Operatoren »Benennen/Analysieren« (10 Bewertungseinheiten), »Interpretieren« der Liedtexte (10 Bewertungseinheiten), »Belegen« (10 Bewertungseinheiten) sowie »Entwerfen und Gestalten« (30 Bewertungseinheiten) gegliedert. Die Urteilsbildung taucht in dieser Gliederung nicht auf. Es wird jedoch unter dem Punkt »Entwerfen und Gestalten« die Anforderung formuliert, »in kreativer und offener Schreibform eine begründete Haltung zum Thema dar[z]ustellen«.²⁸ Damit weicht der Erwartungshorizont erheblich von den Operatordefinitionen der EPA ab. Die Schwerpunktsetzung auf »entwerfen und gestalten«, aber auch die hervorgehobene Rolle der Liedinterpretationen, die 1/6tel der Punkte ausmachen, bleiben dem Prüfling unbekannt.

Im Bereich *Wissen* wird der von der Aufgabe geforderte Anspruch, sich mit der Botschaft Jesu auseinanderzusetzen, im Erwartungshorizont nicht eingeholt. Die Botschaft Jesu wird im Erwartungshorizont mit den Materialien gleichgesetzt, wobei die Botschaft Jesu von den Materialien in unterschiedlicher Weise und mit unterschiedlichen Textintentionen aufgegriffen wird. Durch die Materialien wird ein Fokus gesetzt, der die Bearbeitung leiten kann. Eine umfassende und differenzierte Auseinandersetzung mit der Botschaft Jesu im vorgegebenen Zeitrahmen erscheint aus fachwissenschaftlicher Perspektive als eine

²⁸ Diese Subsumierung der Urteilsbildung unter eine Gestaltungsaufgabe erinnert an den Befund aus Bayern, dass die Erwartungshorizonte Urteilsbildungen nicht aufgreifen, wenn diese in Zusammenhang mit Gestaltungsaufgaben gestellt sind (siehe Kapitel 6.3.1).

deutliche Überforderung der Prüflinge. Eine aus dem Erwartungshorizont abgeleitete Aufgabenstellung müsste passender eine Auseinandersetzung mit den Materialien fordern.

Inhaltlich werden die beiden in der Aufgabe benannten Bibelstellen im Erwartungshorizont durch zwei Stichpunkte erläutert, wobei die Analyse der Bergpredigt auf der Ebene einer Reproduktion verbleibt, da inhaltliche Aussagen ausschließlich benannt werden sollen. Zur Perikope über das Doppelgebot der Liebe wird im Erwartungshorizont der Stichpunkt formuliert: »Gottesliebe und Nächstenliebe als höchste ethisch-moralische Grundsätze für das Zusammenleben deuten« [E-TH, 2014 (2)]. Da die Deutung nicht als optional gekennzeichnet ist, bleibt ihr Verbindlichkeitsgrad uneindeutig. Eine kritische Reflexion der Bibelstelle, die Ableitung christlicher Maßstäbe oder die Einordnung der Perikope in das Gesamtwirken Jesu, wie sie in der geforderten Auseinandersetzung mit der Botschaft Jesu und der Rolle christlicher Maßstäbe für das eigene Leben zu erwarten wären, bleiben aus. Inhaltliche Aspekte, die sich konkret auf die Fragestellung beziehen, fehlen dem Erwartungshorizont.

Eine kritische Distanz lassen auch die inhaltlichen Hinweise zu den Liedtexten (M 2 und M 3) vermissen. Eine Reflexion darüber, inwieweit diese Texte respektive die Textsorte eine Aussage über die Botschaft Jesu und über christliche Maßstäbe treffen können, fehlt.

Ähnliche Leerstellen lassen sich im Bereich *Einstellungen und Haltungen* konstatieren: Dass »eine begründete Haltung zum Thema« und »persönliche[] Erfahrungen situationsgerecht und adressatenbezogen« einbezogen werden sollen, ist zwar erwähnt, wie eine Zustimmung oder Ablehnung der Botschaft Jesu und christlicher Maßstäbe jedoch konkret aussehen könnte und nach welchen Maßstäben die Begründung zu beurteilen ist, wird nicht genannt. Zugunsten der Auswertungsobjektivität wäre eine solche Ausformulierung vor allem in Bezug auf einen kritischen Umgang mit der Botschaft Jesu oder einer Ablehnung dieser von besonderem Interesse, da die Spannung zwischen der Offenheit der Urteilsbildung und der konfessionellen Bindung des Faches an dieser Stelle sichtbar werden kann. Es wäre folglich zu klären, wie damit umzugehen ist. Gleiches gilt in Bezug auf eine Position, welche christliche Maßstäbe für persönliche Entscheidungen als irrelevant erachtet (vgl. zur Bekenntnisorientierung Kapitel 9.3).

Im Bereich *Fähigkeiten und Fertigkeiten* werden zum einen Qualitätsmerkmale benannt, welche sich auf die Form des Essays beziehen: Dieses soll situationsgerecht und adressatenbezogen sein. Da die Aufgabenstellung jedoch weder eine Situation noch mögliche Adressaten vorgibt und ein Essay diese auch nicht notwendigerweise benennt, bleibt unklar, von welcher Situation und von welchen Adressaten der Prüfling beim Schreiben ausgehen soll. Diese Kriterien sind daher nicht prüfbar. Eine kreative und offene Schreibform als zwei weitere genannte Qualitätsmerkmale lassen ohne Maßstab keine Graduierung zu.

In Bezug auf die Urteilsbildung selbst bleibt dem Prüfling die an zwei Stellen genannte Forderung des Erwartungshorizonts, in das Essay persönliche Erfahrungen einzubeziehen, undurchsichtig. Beide Aufgaben – die Auseinandersetzung mit der Botschaft Jesu und die Bewertung der Rolle christlicher Maßstäbe – ließen sich auch ohne den Rückgriff auf persönliche Erfahrungen lösen. Der Einbezug geeigneter biblischer Texte ist durch den Materialbezug bereits abgedeckt, wobei nicht deutlich wird, ob die Prüflinge darüber hinaus noch selbst ausgewählte Bibeltexte anführen sollen. Die Gewichtung dieses Aspekts mit 10 von 60 Bewertungseinheiten lässt darauf schließen, wobei diese Gewichtung – wie bereits wiederholt kritisch angemerkt wurde – den Prüflingen nicht ersichtlich ist. Das Kriterium, dass durch den Essay »in pointierter Form Denkanstöße« geschaffen werden sollen, dient ebenso ausschließlich als ein Impuls – es kann jedoch nicht als ein objektivierendes Analysekriterium des Textes dienen, da es eine hohe Deutungsoffenheit besitzt, die subjektiv durch die korrigierenden Lehrkräfte gefüllt werden muss. Auf den Bereich »Gestalten und Entwerfen« entfallen mit 30 von 60 Punkten die Hälfte aller zu erreichenden Punkte. Zwar klingen in der Ausformulierung auch inhaltliche Aspekte wie eine »begründete Haltung zum Thema« mit an. Sie sind jedoch sekundär. Aspekte eines Urteilsbildungsprozesses sind in diesem Erwartungshorizont in den Hintergrund gedrängt. Die erwarteten Fähigkeiten sind in erster Linie Formulierungs- und Gestaltungskompetenzen, Analysekompetenzen der Materialien und Interpretationskompetenzen in Bezug auf das Gedicht. Die Validität der Aufgabe, das heißt, dass der Test das prüft, was er vorgibt zu prüfen, wird durch den Erwartungshorizont unterlaufen (vgl. zur Validität Kapitel 3.3.2).

Die Analyse des Erwartungshorizontes macht zwei zentrale Probleme desselben deutlich: Zum einen, dass den Prüflingen die Anforderungen nicht transparent gemacht werden. Zum anderen fehlt die inhaltliche Ausformulierung der erwarteten Aspekte, so dass der Erwartungshorizont in sich selbst intransparent bleibt. Um eine höhere Transparenz herzustellen und Auswertungsobjektivität zu sichern, müssten sowohl die Aufgabenstellung als auch der Erwartungshorizont konkretisiert werden.

7.5.2 Exkurs: Unstimmigkeit zwischen Operatoren und Erwartungen

Die bereits im vorangegangenen Beispiel sichtbar gewordenen Unstimmigkeiten zwischen den Operatoren der Aufgabenstellung und den Anforderungen der Erwartungshorizonte soll anhand eines Aufgabensets aus dem Jahr 2016 noch einmal problematisiert werden. Es wurden die Aufgaben gestellt:

»Der Herr hat's gegeben, der Herr hat's genommen; der Name des Herrn sei gelobt.

Die Bibel nach der Übersetzung Martin Luthers, Stuttgart 1999, Hiob 1,21.

Entwerfen Sie eine Trauerrede oder einen Essay zum Thema Sterben und Tod.

Setzen Sie sich dabei mit der Theodizeefrage auseinander und zeigen Sie Perspektiven im Umgang mit Leid auf. Beziehen Sie die Materialien mit ein.«²⁹ [TH, 2016, eA (1)]

»Sich auseinandersetzen mit« und »Perspektiven aufzeigen« signalisieren zwei Anforderungen, die sich durch das Adverb »dabei« auf die Aufforderung »entwerfen« beziehen lassen. Das »Einbeziehen« der Materialien ist keine einzeln gelistete Aufforderung unter den in den EPA definierten Operatoren und ist im Zusammenhang mit dem Entwerfen der Trauerrede oder des Essays zu verstehen. Es ist folglich davon auszugehen, dass ein »Gesamtprodukt« in Form einer Trauerrede oder eines Essays erwartet wird, welches in den 270 Minuten Erarbeitungszeit zu formulieren ist.

Die bereits im vorangegangenen Beispiel dargestellte Unstimmigkeit zwischen den Operatoren der Aufgabenstellung und den Operatoren des Erwartungshorizontes zeigt sich auch hier: Nur einer der vier den Erwartungshorizont gliedernden Operatoren »Darstellen/Analysieren« (AFB I/II; 10 Bewertungseinheiten), »Interpretieren« (AFB III, 10 Bewertungseinheiten), »Erläutern« (AFB II; 10 Bewertungseinheiten) sowie »Entwerfen und Gestalten« (AFB III; 30 Bewertungseinheiten) wird in der Aufgabenstellung erwähnt (entwerfen). Die anderen im Erwartungshorizont benannten Anforderungen sind dem Prüfling nicht ersichtlich.

Die Aufforderung »Darstellen/Analysieren« bezieht sich auf die Materialien, wobei erwähnt ist, dass »die Materialien 1 bis 4 aufgabenbezogen [erschlossen] und wahlweise lösungsorientiert [einbezogen]« werden sollen [E-TH, 2016, eA (1)]. Die gewährte Auswahl (»wahlweise«) trifft in Bezug auf Material 5 – das Gedicht *Hiob* von NELLY SACHS – nicht mehr zu, da hier auf die Forderung der sachgemäßen Analyse und schlüssigen Auslegung des Gedichts »auf Basis der methodisch reflektierten Deutung« 10 Bewertungseinheiten vergeben werden. Wenn der Prüfling sich also entschieden hat, andere Materialien einzubeziehen und das Gedicht außen vor zu lassen, kann er im besten Fall 12 Notenpunkte erreichen, was der Note 1,75 entspricht. In Anbetracht der Tatsache, dass eine Interpretation des Gedichts in der Aufgabenstellung nicht explizit gefordert wurde, kann an diesem Beispiel gezeigt werden, dass die formale Gliederung der Erwartungshorizonte in Operatoren, die in den Aufgabenstellungen nicht adäquat erscheinen, zu Unstimmigkeiten zwischen Anforderungen und Erwartungen führen. Ebenso führt die Intransparenz der Verteilung der Bewertungs-

²⁹ Hervorhebung wie im Original. Weitere Formatierungen wurden nicht übernommen.

einheiten, die ausschließlich in den – den Prüflingen nicht zugänglichen – Erwartungshorizonten stehen, dazu, dass die Prüflinge keine Möglichkeit haben, den Schwerpunktsetzungen der Erwartungshorizonte gerecht zu werden: Im benannten Fall macht die Interpretation des Gedichts rund 17 % und die Auseinandersetzung mit den Materialien insgesamt rund 33 % – also ein Drittel – der Gesamtbewertung aus, obwohl dies die Formulierung der Aufgabenstellung nicht vermuten lässt. Die Validität, welche Auskunft darüber gibt, ob ein Test misst, was er vorgibt zu messen (vgl. Kapitel 3.3.2), ist hier durch die fehlende Passung zwischen Aufgabenstellung und Erwartungshorizont gestört: Die Bewertung richtet sich maßgeblich auf Gestaltungsaspekte (30 Bewertungseinheiten). Die Auseinandersetzung mit der Theodizeefrage und die Perspektiven für einen Umgang mit Leid spielen im Erwartungshorizont nur eine untergeordnete Rolle (10 Bewertungseinheiten). Sie sind mit Blick auf die Bewertungseinheiten mit der Gedichtinterpretation gleichgesetzt, die in der Aufgabenstellung nicht als Anforderung formuliert wird.

7.5.3 Urteilskompetenz im Zusammenhang mit einer Gedichtinterpretation

Ein letztes Aufgabenset, welches aus Thüringen in den Blick genommen werden soll, enthält bei insgesamt drei Aufgabenstellungen zwei Urteilsbildungsaufgaben. Das vollständige Aufgabenset lautet folgendermaßen:

»Beschreiben Sie das Verständnis Martin Luthers zum Thema »Freiheit« entsprechend seiner Schrift »Von der Freyheith eines Christenmenschen« (Material 1).

Erläutern Sie unter Zuhilfenahme mindestens zweier selbstgewählter Beispiele, inwiefern das Freiheitsverständnis Luthers heute noch Relevanz besitzt.

Interpretieren Sie – auch mit Hilfe Ihrer Kenntnisse vom christlichen Menschenbild und der evangelischen Sozialethik – das Gedicht »Stufen« von Hermann Hesse (Material 2) unter der Fragestellung, wie sich Freiheit im Leben verantwortlich gestalten lässt.«³⁰ [TH, 2017 (1)]

Bei beiden Aufgabenstellungen zur Urteilsbildung handelt es sich um in Kapitel 5.2.3 beschriebene Grenzfälle, da die Operatoren ihrer EPA-Definition entsprechend nicht auf eine Urteilsbildung schließen lassen. »Erläutern« ist definiert als »einen Sachverhalt, eine These etc. ggf. mit zusätzlichen Informationen und Beispielen nachvollziehbar veranschaulichen«³¹ und ist dem AFB II zuzuordnen. Die Frage nach der gegenwärtigen Relevanz des Lutherischen Frei-

³⁰ Das Gedicht ist im Anhang zu finden (vgl. Anhang V).

³¹ EPA, 13.

heitsverständnisses kann jedoch nicht »objektiv« in einer Transferleistung beantwortet werden, sondern bedarf durch ihre Kontroversität einer Beurteilung. Die Fragestellung ist folglich dem AFB III zuzuordnen und hätte eine Stellungnahme oder Beurteilung fordern müssen. Im Erwartungshorizont des Aufgabensets ist die Zuordnung zu AFB II erfolgt, es wird darüber hinaus eine eigenständige Positionierung erwartet (siehe unten), die auf inhaltlicher Ebene auch in der Aufgabenstellung verankert ist – der Operator ist folglich nicht passend.

Adäquat verhält es sich mit der sich anschließenden Aufgabe, die den Operator »interpretieren« nutzt, welcher laut seiner Definition in den EPA eine sachgemäße Analyse, ein methodisch reflektiertes Deuten sowie eine schlüssige Gesamtauslegung verlangt (vgl. Kapitel 5.2.3).³² Eine begründete Stellungnahme ist darin nicht vorgesehen und muss aus der enthaltenen Fragestellung »wie sich Freiheit im Leben gestalten lässt« [TH, 2017 (1)] abgeleitet werden. Dass der Prüfling selbst auf die geforderte begründete Stellungnahme schließen muss, die der Erwartungshorizont verlangt (siehe unten), stellt eine unnötige Hürde in der Aufgabenstellung dar und entspricht nicht den EPA, die die Verwendung der passenden Operatoren fordern.

Im Gegensatz zu den im Vorangegangenen betrachteten Beispielen wird der Erwartungshorizont durch die Operatoren strukturiert, die auch in der Aufgabenstellung zu finden sind:³³

<i>Erwartungshorizont</i>	AB ³⁴	BE
Beschreiben: [...]	I	10
Erläutern: Sachverhalte durch Wissen und Einsichten in einen Zusammenhang einordnen und deuten sowie durch Beispiele verdeutlichen	II	20
– strukturierte Erläuterung der Rechtfertigung und christlichen Freiheit – Begriffe: – Sünde/Schuld – Rechtfertigung – Glaube – sola scriptura – sola fide – sola gratia – solus christus – Bezugspunkt für Handeln Gottes: der Mensch in Geschichte und Gegenwart – mindestens zwei Beispiele (pro oder contra) nach eigener Wahl – erkennbarer Gegenwartsbezug		

³² Vgl. a. a. O., 14.

³³ Formatierung wie im Original.

³⁴ Anforderungsbereich [Fußnote wie im Original].

Fortsetzung

<i>Erwartungshorizont</i>	AB	BE
– begründete Positionierung		
Interpretieren: Sinnzusammenhänge erschließen und eine begründete Stellungnahme abgeben	III	30
– anthropologische Grunderkenntnisse:		
– Vergänglichkeit		
– der eigene Wert einer jeden Lebensstufe		
– Bindungsfähigkeit		
– Freiheit zur Entscheidung		
– steter Wandel		
– Fähigkeit fortzuschreiten, sich zu entwickeln (zu transzendieren)		
– Freiheit nicht nur von, sondern auch wozu ...		
– Freiheit bedeutet Verantwortung.		
– Freiheit muss gestaltet werden.		
– Bindungen aus freiem Willen gehören zur Freiheit.		
– Notwendigkeit wertorientierter Entscheidungen		
– Fundamente: Glaube und Vernunft		
– Mensch als Individual- und Sozialwesen hat Verantwortung für sich und andere (Grenzen der Freiheit).		
– eschatologische Dimension		
– Freiheit des Christenmenschen ist Gnade und Entscheidung (Anruf Gottes und Antwort des Menschen).		
– Abschied vom Weltlichen verheißt Heil (Gesundung).		
– Textbezug bei der Argumentation		

[E-TH, 2017 (1)]

Der Fokus der Analyse soll auf der letzten Aufgabenstellung liegen, welche durch ihre außergewöhnliche Verbindung einer Gedichtinterpretation mit einer Urteilsbildung über die Gestaltung von Freiheit im Leben hervorsteht: Auf die Aufgabe entfällt die Hälfte der zu erreichenden Bewertungseinheiten, so dass als Richtwert für die Bearbeitungszeit 135 Minuten angenommen werden können.

Auf inhaltlicher Ebene wird im Erwartungshorizont die Deutungsoffenheit, welche aufgrund der Aufgabenstellung vorauszusetzen wäre, nicht offenkundig. Konträre Positionen zur Fragestellung sind nicht benannt. Die Aussagen, welche in Bezug auf Freiheit getroffen werden, beispielsweise dass Bindungen aus freiem Willen zur Freiheit gehören, werden nicht kritisch hinterfragt oder in ihrer Positionalität reflektiert, sondern erscheinen als eine objektive Lösung der Aufgabenstellung. Darüber hinaus enthalten die getroffenen Aussagen normative Setzungen wie beispielsweise »Notwendigkeit wertorientierter Entscheidungen«. Die im Erwartungshorizont genannten Lösungen lassen sich nicht unmittelbar

aus dem Gedicht HERMANN HESSES ableiten und werden nicht mit diesem in Verbindung gebracht. So erschließt sich beispielsweise nicht, wie sich eine eschatologische Dimension in die Interpretation und in den Argumentationsgang fügen könnte und ob sie zur Lösung der Aufgabenstellung essenziell ist oder nicht. Eine höhere Anschaulichkeit des Erwartungshorizontes könnte erreicht werden, wenn Verse des Gedichts mit Deutungen in Verbindung gebracht würden.

Die dem Erwartungshorizont auf inhaltlicher Ebene fehlende kritische Distanz zu den vorgeschlagenen Deutungsansätzen wird für den Bereich *Einstellungen und Haltungen* nochmals relevant: Persönliche Einstellungen und Haltungen geraten im Erwartungshorizont nicht in den Blick.

Die für die geforderte Erschließung der Sinnzusammenhänge und begründete Stellungnahme benötigten *Fähigkeiten und Fertigkeiten* sowie ihre Qualitätskriterien bleiben ebenso unbenannt. Es gibt lediglich einen Hinweis darauf, dass ein Textbezug in der Argumentation herzustellen ist.

7.6 Zusammenfassung der fallorientierten Ergebnisse

»Will man Auswertungsobjektivität in Zentralen Prüfungen erreichen, dann kommt alles auf die Korrekturanleitungen an. Korrekturanleitungen geben darüber Auskunft, welche Leistungen erbracht werden müssen, um Aufgaben vollständig, teilweise oder in Ansätzen zu lösen. Folglich ist die gute Korrekturanweisung, die übereinstimmende Urteile gestattet, erschöpfend und detailliert, enthält echte Schülerantwortbeispiele und darüber hinaus eine von den Beispielen abstrahierte Charakterisierung der Antwortidee.«³⁵

Die Deutschdidaktikerin JULIANE KÖSTER formuliert im Zitat einen »Optimalanspruch« an Erwartungshorizonte, dem die Erwartungshorizonte für Urteilsbildungsaufgaben im Untersuchungszeitraum kaum genügen. Was durch die kategorienorientierte Auswertung in Kapitel 6 nicht erfasst, anhand der vertiefenden Einzelfallanalysen jedoch deutlich wurde, ist, dass Qualitätsmerkmale für Urteilsbildungen in den Erwartungshorizonten mitunter völlig fehlen oder aber auch unzureichend auf die konkrete Fragestellung hin ausdifferenziert werden. Es ist darüber hinaus deutlich geworden, dass die für Urteilsbildungsaufgaben vorgesehenen Bearbeitungszeiten mitunter stark variieren. Während in Bezug auf die Erwartungshorizonte aus Bayern und Niedersachsen die Auswertungsobjektivität als ein zentrales Problem benannt werden kann, welches dadurch entsteht, dass es für eine graduelle Bewertung an Niveaudifferenzierungen mangelt, ist im Hinblick auf aus Thüringen stammende Erwartungshorizonte vor

³⁵ KÖSTER, Leistung und Qualität, 31 f.

allen die Frage nach der Validität zentral. Abstriche in Bezug auf die Validität der Aufgabenstellungen entstehen durch die dargestellten Unstimmigkeiten in der Verwendung der Operatoren. Diese führen zu mangelnder Transparenz, was sich sowohl negativ für die Prüflinge auswirkt, da sie die Schwerpunktsetzungen der Erwartungshorizonte nicht kennen, als auch negativ für die korrigierenden Lehrkräfte, da sie mit den Unstimmigkeiten umgehen und Lösungen finden müssen.

Die Ergebnisse der Fallanalysen fließen in die im folgenden Dargestellten Profilbildungen der drei untersuchten Bundesländer ein, so dass Kapitel 8 als eine Zusammenfassung der Kapitel 6 und 7 betrachtet werden kann.

8 Aufgabenkulturen und Bewertungsvorgaben – Profilierung der Bundesländer

Die Analyse der Aufgabenstellungen, in denen ein Urteil über eine ethische und/oder religiöse Fragestellung zu fällen ist, hat länderspezifische Prüfungskonventionen sowohl hinsichtlich der Aufgabengestaltung als auch in Bezug auf die Erwartungshorizonte ergeben, die im Folgenden interpretiert werden. Dazu werden in diesem Kapitel zunächst Profile erstellt, welche die zentralen Charakteristika der drei untersuchten Bundesländer in den Fokus rücken.¹ Die Profilbildungen verstehen sich als Interpretationen der Analyseergebnisse vor dem Hintergrund der zentralen Forschungsfragen der Studie:

Wie wird Urteilsbildung im Kontext schriftlicher Abiturprüfungen getestet (Häufigkeit, Themen, Reflexionskriterien etc.)?

Welche Bewertungskriterien werden in den Erwartungshorizonten benannt (Qualitätsmerkmale, Umgang mit Einstellungen und Haltungen, Definition konträrer Positionen etc.)?

8.1 Bayern – Christliche Weltsicht als Reflexionsmedium ethischer Fragestellungen

8.1.1 Aufgabenstellungen zu Urteilsfähigkeit aus Bayern

Für das bayerische Abitur im Fach Evangelische Religionslehre ist die hohe Anzahl an Aufgabenstellungen je Aufgabenset charakteristisch, wodurch die Bedeutung eigenständiger Urteilsbildungsaufgaben in ethischen und religiösen Fragestellungen in Bezug auf die Gesamtaufgabenzahl prozentual gering ist (~ 7%). Dennoch kommen Urteilsbildungsaufgaben in knapp der Hälfte aller

¹ Methodisch entsprechen die Charakterisierungen keinen Typenbildungen, da nicht gleiche Merkmale zur Konstruktion der Typologien führen, sondern die Herkunft der Fälle aus einem Bundesland das Kriterium der Gruppierung darstellen. Es handelt sich folglich um »natürliche Gruppen« (vgl. KUCKARTZ, Qualitative Inhaltsanalyse, 156).

Aufgabensets vor (~ 42 %). Die Vielzahl an Aufgabenstellungen wirkt sich auch auf die ohnehin im Vergleich zu den beiden anderen Bundesländern niedrige Bearbeitungszeit der Aufgaben aus. Die Rahmenbedingungen bedingen folglich, dass die eigenständige Urteilsbildung im schriftlichen Abitur in Bayern im Vergleich zu den anderen beiden untersuchten Bundesländern einen untergeordneten Stellenwert besitzt. Die Seltenheit eigenständiger Urteilsbildungsaufgaben wird dadurch ausgeglichen, dass andere, aus der vorliegenden Studie ausgeschlossene Aufgabenstellungen aus Anforderungsbereich (AFB) III in Bayern eine größere Rolle spielen wie beispielsweise Urteilsbildungen mit Perspektivwechsel oder Aufgabenstellungen mit dem Operator »Perspektiven entwickeln«.² Die im bayerischen Abitur des Untersuchungszeitraums zu beurteilenden Fragestellungen sind überwiegend ethische Fragen ohne dezidiert religiösen Bezug, welche vor dem Hintergrund christlicher Sichtweisen reflektiert werden sollen. Eine typische Formulierung für die Reflexionsaufforderung ist beispielsweise: »[...] berücksichtigen Sie dabei evangelische Einsichten zu Gott und Mensch« [BY, 2014 (4)]. Durch Formulierungen dieser Art wird die in den Einheitlichen Prüfungsanforderungen (EPA) formulierte Anforderung, die »[...] eigene[] Urteilsbildung unter Beachtung biblischer, theologischer und ethischer Kategorien«³ zu reflektieren, aufgegriffen. Den Aufgabenstellungen liegen häufig Zeitungsartikel und/oder theologische Sachtexte zugrunde.

8.1.2 Erwartungshorizonte zu Urteilsaufgaben aus Bayern

Die Erwartungshorizonte für Urteilsbildungsaufgaben bleiben im bayerischen Abitur in mehreren Fällen auf einer allgemeinen Ebene, indem sie eine kritische Stellung oder eine eigene Antwort fordern, ohne dass inhaltliche Aspekte angesprochen werden. Besonders auffällig ist, dass Urteilsbildungen, welche in Zusammenhang mit Gestaltungsaufgaben stehen (beispielsweise, wenn ein Leserbrief oder eine Schülerzeitungsartikel verfasst werden soll), in den Erwartungshorizonten nicht aufgegriffen werden. In diesen Fällen werden im Erwartungshorizont Qualitätsmerkmale der Gestaltung benannt wie beispielsweise »Adressaten- und Situationsbezug« oder eine verständliche Sprache, jedoch keine Qualitätsmerkmale für die geforderte Urteilsbildung.

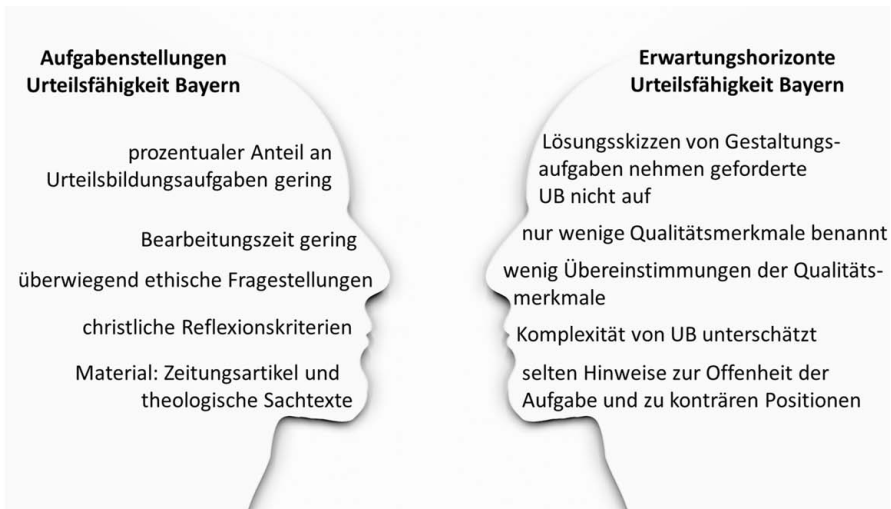
Insgesamt werden Qualitätsmerkmale für Urteilsfähigkeit in den Erwartungshorizonten nur selten benannt, was neben der geringen Bearbeitungszeit für diese Aufgaben einen Hinweis darauf gibt, dass die Komplexität der Anforderungen, die Urteilsbildungsaufgaben an die Prüflinge stellen, unterschätzt

² Vgl. zum Ausschluss dieser Aufgabenstellungen aus dem Untersuchungskorpus Kapitel 5.2.2.

³ EPA, 12.

wird. Die in Bezug auf Urteilsbildung in den Erwartungshorizonten genannten Qualitätsmerkmale, die nicht explizit von der Aufgabe gefordert wurden, weisen darüber hinaus kaum Übereinstimmungen auf: Neben dem Urteilsentscheid selbst, werden Differenziertheit und die Begründung in vier Fällen als häufigste Übereinstimmungen genannt. Hinweise auf die Offenheit des Urteilsentscheides existieren, sie sind aber nicht die Regel (in zwei von elf Fällen). Ebenso verhält es sich in Bezug auf die Benennung konträrer Positionen, die in drei von elf Fällen zu finden ist.

Die folgende Grafik visualisiert die herausgearbeitete Aufgabenkultur und Merkmale der Erwartungshorizonte der in Bayern gestellten Urteilsbildungsaufgaben in Form zweier Gesichtsprofile, wobei explizit erwähnt sei, dass die Gegenüberstellung nur einer Systematisierung der Bereiche dient und nicht als Inkohärenz der Bereiche zu interpretieren ist.



* UB = Urteilsbildung

Abb. 19: Charakteristika der Aufgabenkultur und Bewertungsvorgaben von Urteilsbildungsaufgaben aus Bayern.

8.2 Niedersachsen – Durch Theologinnen und Theologen vermittelte Weltsicht in formalisiertem Rahmen

8.2.1 Aufgabenstellungen zu Urteilsfähigkeit aus Niedersachsen

Die Aufgabenkultur in Bezug auf Urteilsbildung lässt sich in den aus Niedersachsen stammenden Aufgabensets ebenso wie in den dazugehörigen Erwar-

tungshorizonten als formalisiert beschreiben. Die Aufgabensets haben stets die gleiche Anzahl an Aufgabenstellungen, welche sich eindeutig den AFB zuordnen lassen, einer übergeordneten Thematik unterstellt sind und hinsichtlich ihres Schwierigkeitsgrades einem kumulativen Aufbau folgen. Den Aufgabensets liegt in der Regel *ein* Material zugrunde. Dieses besteht überwiegend aus einem theologischen Sachtext. Diese immer wiederkehrenden Rahmenbedingungen erzeugen Erwartungssicherheit und Transparenz hinsichtlich der Anforderung. Die aufeinander aufbauenden Aufgabenstellungen entlasten die Urteilsbildung, da in der Regel bereits durch die vorangehenden Aufgabenstellungen Vorüberlegungen getroffen wurden. Gleichzeitig sind die in Zusammenhang mit Urteilsbildungen gestellten Gestaltungsaufgaben mit nur 25 % in Niedersachsen ebenso wie die Diversität der Textsorten des zugrunde gelegten Materials niedriger als in Thüringen und Bayern.

In den großzügigen Bearbeitungszeiten und der Häufigkeit der Urteilsbildungsaufgaben wird deutlich, dass die Verfasserinnen und Verfasser der Abituraufgabenstellungen der eigenständigen Urteilsbildung einen zentralen Stellenwert beimessen: In 75 % aller im Untersuchungszeitraum in Niedersachsen gestellten Aufgabensets wurde eine eigenständige Urteilsbildung zu einer ethischen und/oder religiösen Fragestellung gefordert. Der Schwerpunkt dieser Aufgaben liegt auf der in den EPA formulierten Anforderung: »Erörtern von fachspezifischen Positionen, Thesen und Problemen mit dem Ziel einer begründeten und überzeugenden Stellungnahmen«⁴, während der Anspruch des »Reflektieren[s] der eigenen Urteilsbildung unter Beachtung biblischer, theologischer und ethischer Kategorien«⁵, der in Aufgabenstellungen zur Urteilsbildung aus Bayern wiederholt gefordert wird, in den Hintergrund rückt. Die fachspezifischen Positionen, über welche geurteilt werden soll, sind in der Regel christlich und beziehen sich sowohl auf religiöse als auch auf Fragestellungen ohne dezidiert religiösen Bezug. Durch die Aufforderung, eine spezifische christliche Positionierung zu bewerten, liegt die theologische Angemessenheit als eine Beurteilungskategorie nahe, die nicht in den Fokus rückt, wenn christliche Sichtweisen als Reflexionskategorie der eigenen Urteilsbildung auftreten, wie es in den bayerischen Aufgabenstellungen häufig der Fall ist. Darüber hinaus wird die theologische Argumentationsfähigkeit der Schülerinnen und Schüler durch die Auseinandersetzung mit theologischen Positionierungen stark gefordert. Zudem wird die Pluralität innerhalb der christlichen Perspektive durch diese Art der Aufgabenstellungen besonders deutlich, da von vornherein in Frage gestellt wird, dass die spezifische Perspektive einen generellen Gültigkeitsanspruch erheben kann.

⁴ Vgl. fachspezifische Beschreibung der AFB in den EPA, 12.

⁵ Vgl. ebd.

Dass in Bezug auf Urteilsbildung keine Schwierigkeitsdifferenzierung zwischen erhöhtem und grundlegendem Anforderungsniveau stattfindet, legt auch vor dem Hintergrund der generell sehr allgemein formulierten Qualitätsmerkmale in diesem Bereich nahe, dass kein ausdifferenziertes Konzept von Urteilsfähigkeit besteht, welches eine Benennung unterschiedlicher Niveaustufen und damit eine graduelle Bewertung möglich macht.

8.2.2 Erwartungshorizonte zu Urteilsaufgaben aus Niedersachsen

Die aus Niedersachsen stammenden Erwartungshorizonte sind im Vergleich zu den beiden anderen untersuchten Bundesländern die umfangreichsten. Sie setzen zudem die Vorgaben der EPA am gewissenhaftesten um. Alle Erwartungshorizonte der zwanzig Aufgabenstellungen, die eine eigenständige Urteilsfähigkeit in ethischen und religiösen Fragen im Untersuchungszeitraum fordern, gehen im Erwartungshorizont auf die Urteilsbildung ein. Jedoch gibt es auch in den aus Niedersachsen stammenden Erwartungshorizonten keine Qualitätsmerkmale, welche durchgängig im Zusammenhang mit Urteilsbildung auftauchen. Die höchste Übereinstimmung liegt in 13 Fällen vor, in denen »Eigenständigkeit« und das Geben einer Begründung als Qualitätsmerkmale benannt sind. Der quantitative Umfang der von den Prüflingen anzufertigenden Urteilsentscheide wird in zehn der 20 Aufgabenstellungen explizit als Qualitätsmerkmal nivelliert. Ein weiteres Herausstellungsmerkmal niedersächsischer Erwartungshorizonte ist, dass die Offenheit der Urteilsbildung in 14 von 20 Fällen explizit angesprochen wird. Dies entspricht auch einem Hinweis aus dem niedersächsischen Lehrplan für die gymnasiale Oberstufe, welcher unter dem Kapitel »Leistungsfeststellung und Leistungsbewertung« den Hinweis gibt, dass es bei der Bewertung um »überprüfbare Qualifikationen des Wissens, Argumentierens und gestalterischen Handelns geht, nicht aber um religiöse, politische oder moralische Einstellungen der Schülerinnen und Schüler.«⁶ Ebenso bemerkenswert ist, dass in 80 % der Erwartungshorizonte konträre Positionen zu den Fragestellungen inhaltlich ausformuliert werden.

Auch hier sollen die herausgearbeiteten Merkmale der Prüfungskonventionen in Bezug auf Urteilsbildungsaufgaben grafisch in zwei Profilen zusammenfassend visualisiert werden:

⁶ NIEDERSÄCHSISCHES KULTUSMINISTERIUM (Hrsg.), Kerncurriculum, 31.



* UB = Urteilsbildung

Abb. 20: Charakteristika der Aufgabenkultur und Bewertungsvorgaben von Urteilsbildungsaufgaben aus Niedersachsen.

8.3 Thüringen – Komplexe Aufgabenstellungen unter Einbezug biblischer Texte

8.3.1 Aufgabenstellungen zu Urteilsfähigkeit aus Thüringen

Die Abnahme schriftlicher Abiturprüfungen im Fach Evangelische Religionslehre ist im von Säkularisierung geprägtem Thüringen mit Zahlen zwischen acht und 19 Prüflingen pro Jahr sehr selten, was jedoch nicht zum Ausschluss aus der Untersuchung geführt hat (vgl. Kapitel 5.1).

Besonders auffällig ist an den aus Thüringen stammenden Abituraufgabensets, dass die Handlungsaufforderungen meist in ein bis zwei Sätzen zu einer Gesamtaufgabenstellung miteinander verknüpft sind. Auch wenn die Aufgabensets aus mehreren Sätzen bestehen, gibt es keine Strukturierung der Aufgabe durch numerische Gliederungspunkte. Ein kumulativer Aufbau des Schwierigkeitsgrades der Aufgabenstellungen, wie es in den Aufgabensets aus Niedersachsen konsequent der Fall ist, lässt sich in den Aufgabensets aus Thüringen nicht erkennen. Die dadurch entstehende Komplexität wird in einigen Fällen durch sehr offen formulierte Fragestellungen zusätzlich erhöht. Einzigartig im Untersuchungszeitraum und für Thüringen markant ist ein Aufgabenset, welches aus nur einer Handlungsaufforderung besteht.⁷ In dieser sollen die Prüflinge

⁷ Unbeachtet bleibt damit die Vorgabe der EPA, dass sich »[e]ine Prüfungsaufgabe für die schriftliche Abiturprüfung [...] aus mehreren Teilaufgaben zusammen[setzt]« (EPA, 18).

unter Einbezug der Materialien, geeigneter Bibeltex te und eigener Erfahrungen eine Rede in Auseinandersetzung mit einem Zitat der evangelischen Theologin MARGOT KÄBMANNs verfassen [TH, 2014 (1)].

Urteilsbildung wird im schriftlichen Abitur in Thüringen überwiegend an religiösen Fragestellungen geprüft, die in fünf von elf Fällen auch mit christlichen Reflexionsanforderungen verknüpft sind. Die in Thüringen gestellten Urteilsbildungsaufgaben stechen durch die Vielzahl an lyrischen Texten hervor, welche die Prüflinge in die Urteilsbildung einbeziehen oder beurteilen sollen (beispielsweise Liedtexte von LUDWIG VON ZINZENDORF, den TOTEN HOSEN und REINHARD MEY oder das Gedicht *Hiob* von NELLY SACHS). Ebenso ist die explizite Forderung des Einbezugs von Bibelversen und -perikopen im Zusammenhang mit Urteilsbildung ein Alleinstellungsmerkmal Thüringens.

8.3.2 Erwartungshorizonte zu Urteilsbildungsaufgaben aus Thüringen

Ebenso wie in den Erwartungshorizonten aus Niedersachsen und Bayern gibt es in den Erwartungshorizonten aus Thüringen neben dem Urteilsentscheid selbst nur wenige, sehr allgemein formuliert Qualitätsmerkmale, die in mehreren Urteilsbildungsaufgaben übereinstimmen. Die meisten Übereinstimmungen bestehen in der Forderung nach einer Begründung (in neun von elf Fällen), nach einem Aufgabenbezug (sechs von elf) und darin, dass die Lösung »sachgerecht« sein soll (fünf von elf). Ein ausdifferenziertes Konzept von Urteilsfähigkeit, welches eine Niveaustufung ermöglicht, ist auch in den Erwartungshorizonten aus Thüringen nicht zu finden. Explizite Hinweise zum Umgang mit den Schülermeinungen im Kontext von Urteilsbildungsaufgaben werden nicht gegeben. Auch durch die inhaltlichen Angaben der Lösungsskizzen wird die Offenheit der Aufgaben in vielen Fällen nicht hinreichend deutlich: So gibt beispielsweise der Erwartungshorizont einer Aufgabe, in der ein Essay zur Theodizeefrage verfasst werden soll, eine Deutung eines einzubeziehenden Bibelverses vor, ohne deutlich zu machen, ob dieser Interpretation zwangsläufig gefolgt werden muss, oder nicht: Dort heißt es zu Ps 27,1: »den zeitlos Hoffnung stiftenden Charakter des Psalmverses herausstellen« und »Gott als Kraftquelle zur Orientierung in konfliktreichen Alltagssituationen aufzeigen« [E-TH 2016, eA (1)]. Hinzu kommt – und auch dies wird an genanntem Beispiel deutlich –, dass die inhaltlichen Aussagen in den Erwartungshorizonten in ihrer Positionalität nicht kritisch reflektiert werden. Konträre Positionen zu den Fragestellungen sind in keinem der Erwartungshorizonte inhaltlich benannt.

Ebenso ist kritisch anzumerken, dass die Erwartungshorizonte durch Operatoren strukturiert werden, welche sich in mehreren Fällen in den Aufgabenstellungen nicht wiederfinden lassen (vgl. Kapitel 7.5). Die in den Lösungsskizzen

formulierten Anforderungen sind daher für den Prüfling nicht transparent. Die Erwartungshorizonte unterlaufen somit die Validität der Aufgabenstellung.⁸ Darüber hinaus werden die Operatoren mitunter frei ausgedeutet, was zusätzlich zu Intransparenz der Anforderungen führt. Diese Beobachtungen zusammenfassend lässt sich sagen, dass die Orientierung an den Vorgaben der EPA sehr gering ausfällt. Die in den Erwartungshorizonten vorgenommene Verteilung der Bewertungseinheiten auf die einzelnen Handlungsaufforderungen wird den Prüflingen nicht genannt, so dass diesen die Schwerpunktsetzungen nicht bekannt sind. Die benannten Punkte führen dazu, dass die formale und inhaltliche Ausgestaltung der Erwartungshorizonte in Thüringen sowohl für die Prüflinge als auch für die korrigierenden Lehrpersonen geringe Orientierungskraft aufweisen.

Folgende Grafik soll die herausgearbeiteten »Profillinien« der Aufgabenkultur sowie Charakteristika der aus Thüringen stammenden Erwartungshorizonte zusammenfassen:



* UB = Urteilsbildung

Abb. 21: Charakteristika der Aufgabenkultur und Bewertungsvorgaben von Urteilsbildungsaufgaben aus Thüringen.

Aus den Profilen lassen sich drei zentrale Aspekte ableiten, welche die Ergebnisdiskussion im Folgenden strukturieren sollen: Zum einen zeigt sich, dass sich

⁸ Das Gütekriterium der Validität gibt an, ob ein Test das misst, was er zu messen vorgibt. Im Idealfall werden von einem Test nur die Fähigkeiten erfasst, die diagnostiziert werden sollen, ohne dass das Ergebnis durch andere Wissensbestände oder Fähigkeiten beeinflusst wird (vgl. Kapitel 3.3.2).

der Orientierungsgrad an den Vorgaben der EPA sowohl auf die Transparenz der Anforderungen als auch auf die Auswertungsobjektivität auswirkt. In Kapitel 9.1 wird daher zu diskutieren sein, was die Ergebnisse der Studie über die Passung zwischen den EPA und der tatsächlichen Prüfungsgestaltung aussagen und wie dies in den Kontext der *Educational Governance Forschung* eingeordnet werden kann. Gerade im Kontext der religionspädagogischen Debatte über Möglichkeiten und Grenzen der Standardisierung religiöser Kompetenz ist die Frage nach der Verbindlichkeit und Umsetzung bildungsadministrativer Vorgaben ein zentrales Diskursfeld (zur religionspädagogischen Debatte vgl. Kapitel 2.3).

Ein zweiter zentraler Punkt bezieht sich auf die Gestaltung der Erwartungshorizonte und die für Urteilsfähigkeit benannten Qualitätsmerkmale. In der Profilbildung wurde bereits angedeutet, dass die genannten Qualitätsmerkmale häufig auf allgemeiner Ebene bleiben und keine Niveaustufendifferenzierungen existieren. Dieser Befund soll in Kapitel 9.2 aufgegriffen und im Kontext der Diagnostikforschung diskutiert werden (vgl. zur Diagnostikforschung auch Kapitel 3.3).

Abschließend wird – als ein dritter zentraler Punkt – der Umgang mit der Bekenntnisorientierung des Faches in den Abituraufgaben und Erwartungshorizonten vor dem Hintergrund der Offenheit der Urteilsbildung diskutiert (Kapitel 9.3).

9 Einheit in Vielfalt?

Eine Diskussion der Ergebnisse im Spiegel des Anspruchs auf Vergleichbarkeit

Mit der Einführung der Einheitlichen Prüfungsanforderungen (EPA) ging der Anspruch einher, durch allgemeinverbindliche Standards »die Transparenz, Vergleichbarkeit und Einheitlichkeit der Prüfungsverfahren und -anforderungen weithin zu gewährleisten«.¹ Betrachtet man die Prüfungskonventionen und Bewertungsvorgaben der drei untersuchten Bundesländer, so lassen sich – wie die vorangegangene Zusammenfassung zeigt – länderspezifische Aufgaben- und Prüfungskulturen erkennen, die zu der Frage führen, wie die Differenzen zwischen den Bundesländern einzuschätzen sind. Diese Frage soll im Zentrum der Ergebnisdiskussion stehen.

KÜHN spricht in Bezug auf die EPA von einer »Hybridbildung« alter und neuer Steuerungselemente, da sie mit ihren Vorgaben zur Prüfungsgestaltung in erster Linie Elemente der Inputsteuerung enthalten, gleichzeitig jedoch auch durch Zielvorgaben den Outcome beeinflussen wollen.² Die beiden ersten Kapitel der Ergebnisdiskussion sind an dieser Unterscheidung KÜHNS orientiert: Zunächst sollen die Differenzen zwischen den Bundesländern hinsichtlich der Prüfungsgestaltung in den Blick genommen werden und vor dem Hintergrund der inputsteuernden, gestaltungsbezogenen Elemente der EPA diskutiert werden (vgl. Kapitel 9.1). Dabei wird zwischen länderspezifischen Differenzen zu unterscheiden sein, die kontextsensibel Unterschieden zwischen den Bundesländern gerecht werden (vgl. Kapitel 9.1.1) und jenen, die die Vergleichbarkeit der Anforderungen tendenziell unterminieren (vgl. Kapitel 9.1.2). Die Gestaltungsaspekte der Prüfungsaufgaben sind nicht unmittelbar auf Urteilsfähigkeit bezogen. Die Urteilsbildung bleibt von ihnen jedoch nicht unberührt: Da Gestaltungsaspekte komplexe Aufgaben, wie sie Urteilsbildungen darstellen, wesentlich be-

¹ STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER (Hrsg.), Vereinbarung über Einheitliche Prüfungsanforderungen in der Abiturprüfung. Beschluss der Kultusministerkonferenz vom 01.06.1979 in der Fassung vom 24.10.2008.

² Vgl. KÜHN, Steuerung und Innovation, 309.

dingen, können sie Hürden schaffen oder zur Reduktion von Komplexität beitragen.

Kapitel 9.2 fokussiert vor dem Hintergrund der EPA-Zielvorgaben, die den Outcome betreffen, jene Ergebnisse der Studie, welche sich auf die Qualitätsmerkmale von Urteilsbildung beziehen. Dabei werden die von den Erwartungshorizonten für Urteilsbildung benannten Qualitätsmerkmale diskutiert (vgl. Kapitel 9.2.1) und in den Kontext der im Forschungsdiskurs akzentuierten Qualitätsmerkmale gestellt (vgl. Kapitel 9.2.2). Es schließen sich in Kapitel 9.1.3 Überlegungen an, wie die Ergebnisse der Studie vor dem Hintergrund der Forderungen der pädagogischen Messtheorie zu deuten sind.

Kapitel 9.3 diskutiert abschließend die Rolle der Bekenntnisorientierung für Urteilsbildungen im Fach Religionslehre und stellt die Spezifika von Urteilsbildungen im religionsunterrichtlichen Kontext heraus, welche zum einen in den Themen der Fragestellungen (vgl. Kapitel 9.3.1) und zum anderen in den Reflexionskriterien liegen, an denen Urteilsbildungen gemessen werden (vgl. Kap. 9.3.2).

9.1 Aspekte der Prüfungsgestaltung – Kontextbedingte Differenzen vs. kontextunabhängige Differenzen

Wie sind die bundeslandspezifischen Gestaltungsaspekte der Abiturprüfungen zu bewerten, welche im vorangegangenen Kapitel zusammenfassend dargestellt wurden? Eine pauschale Antwort auf diese Frage würde zu kurz greifen: Zum einen gibt es Differenzen, welche auf die unterschiedlichen Kontexte der Bundesländer reagieren und somit zu wahren sind, da sie kontextsensibel auf regionale Eigenheiten eingehen, anstatt eine standardisierte »Einheitsbildung« zu verfolgen: Die föderale Organisation des Bildungswesens kann aus dieser Perspektive als ein Gewinn gesehen werden, da sie den Bundesländern ermöglicht, ihre Abschlussprüfungen im Fach Religionslehre an ihren religionssoziologischen Kontext anzupassen. Zum anderen gibt es jedoch auch Differenzen, die zu mangelnder Vergleichbarkeit zwischen den Bundesländern und zu Fehlentwicklungen in Hinblick auf die Auswertungsobjektivität führen.³ In diesen Fällen ist ein gerechtes und transparentes Prüfungsverfahren anzustreben, welches zum Teil durch eine engere Umsetzung der EPA-Vorgaben, zum anderen aber auch durch konkretere Vorgaben seitens der EPA gewährleistet werden kann.

³ Vgl. zum System zentraler Abschlussprüfungen und zu Testgütekriterien Kapitel 3.

9.1.1 Differenzen als Chance einer kontextsensiblen Prüfung

Als Merkmale, die auf die religiöse Diversität innerhalb Deutschlands reagieren, können beispielsweise der Religionsbezug der Aufgabenstellung, die Reflexionskriterien oder die Textsorten der Materialien genannt werden. Die in der Studie untersuchten Bundesländer stehen in sehr unterschiedlichen religionssoziologischen Kontexten (vgl. dazu Kapitel 5.1): Thüringen ist überwiegend säkular geprägt, während evangelische Christinnen und Christen in Bayern zwar im Vergleich zu katholischen Christen auch in der Minderheit sind, Gottesgläubigkeit an sich jedoch nichts Außergewöhnliches ist. Ebenso ist in Niedersachsen die Zugehörigkeit zu einer Glaubensgemeinschaft nichts Außergewöhnliches – evangelische Christinnen und Christen sind hier sogar in der Majorität: Rund die Hälfte der Einwohner sind dem evangelischen Glauben zugehörig, während nur 25 % der Niedersachsen konfessionslos sind.⁴

Die unterschiedlichen Schwerpunktsetzungen der Themen können in den unterschiedlichen religionssoziologischen Kontexten eine Erklärung finden. In Thüringen ist ein stärkerer Traditionsbezug erkennbar: Die Bezugnahme auf Bibelverse und -perikopen ist in Thüringen deutlich sichtbar, während diese in den anderen Bundesländern fehlt. Die Aufgabenstellungen fordern mitunter, dass die Schülerinnen und Schüler eigenständig passende Bibelstellen zitieren, um ihre Argumentation zu stützen. Darüber hinaus tritt der Bekenntnisbezug in den Fragestellungen Thüringens am deutlichsten zu Tage. Beispiele bieten die Auseinandersetzung mit der Botschaft Jesu, die Rolle christlicher Werte für eigene Entscheidungen, die Frage: »Lässt sich mit der Bergpredigt Politik machen?«, die Beurteilung der heutigen Relevanz des lutherischen Freiheitsverständnisses oder die Frage nach dem Wert von Gottesbeweisen für den Gottesglauben. Die stärkere Fokussierung auf Bibel und Bekenntnis könnten darin wurzeln, dass sich der Religionsunterricht in einem säkularisierten Umfeld in der Verantwortung sieht, grundlegende religiöse, bibelorientierte Bildung zu fördern. Ein Blick in die aktuellen Lehrpläne anderer Fächer zeigt beispielhaft, dass Religion im allgemeinen Schulwesen in Thüringen nur randständig vorkommt.⁵

»So finden in den Lehrplänen für die Fächer Deutsch, Geografie, Sozialkunde, Englisch und Musik der Klassenstufen 5 bis 12 die für die christliche und auch jüdische und islamische Religion zentralen Begriffe wie Religion, Christentum, Judentum,

⁴ Zum letzten Zensus 2011 waren 51,5 % der Niedersachsen Mitglied in der evangelischen Kirche (vgl. BERND SCHRÖDER, Religion unterrichten in Niedersachsen, in: MARTIN ROTHGANGEL/BERND SCHRÖDER [Hrsg.], Religionsunterricht in den Ländern der Bundesrepublik Deutschland. Neue empirische Daten – Kontexte – Aktuelle Entwicklungen. Leipzig 2020, 239–267, hier 240).

⁵ WERMKE/WIDL, Religion unterrichten in Thüringen, 425.

Islam, Kirche, Synagoge, Moschee oder Antisemitismus keine Erwähnung. In anderen Lehrplänen wird Religion in substantieller Weise als Religionsgemeinschaft oder Konfession verstanden, die ein (konfliktträchtiges) historisches Phänomen darstellt (LP Latein, LP Geschichte) und deren Gestaltungsformen kulturell bedingt sind (LP Französisch, LP Kunst). Der religiöse Glaube ist unwissenschaftlich und irrational (LP Biologie); Religion als »Lebensstil« findet außerhalb der Lehrpläne für die Evangelische, Katholische und Jüdische Religionslehre keinerlei Aufmerksamkeit, und eingeschränkt gilt dies sogar für das Fach Ethik.«⁶

Es ist daher nur folgerichtig, wenn der Religionsunterricht in Thüringen auf diesen Kontext reagiert und auch die Freiheiten zugesprochen bekommt, darauf reagieren zu können. Die Schwerpunktsetzung im bayerischen Abitur auf ethische Fragestellungen der Urteilsbildung, die durch christliche Kategorien reflektiert werden sollen, betonen hingegen – so eine Interpretationsmöglichkeit – die Relevanz christlicher Sichtweisen auf gesellschaftsrelevante Fragestellungen, was in einen Zusammenhang damit gestellt werden kann, dass Religion in Bayern stärker als »Lebensstil« im Alltag verankert ist, als es in Thüringen der Fall ist. Und auch die in Niedersachsen auffällige Form, Urteilsbildungen mit prominenten theologischen Positionen zu verknüpfen, fügt sich gut in den religiös geprägten Kontext dieses Bundeslandes ein. Durch die Selbstbestimmung von Aufgabekulturen kann jedes Bundesland auf die ihm eigenen Kontexte eingehen.

Ein diese Annahme ergänzendes Erklärungsmodell für die unterschiedlichen Schwerpunktsetzungen liegt im Grundverständnis von Religionspädagogik seitens der Verfasserinnen und Verfasser der Abituraufgabenstellungen. So sind im religionspädagogischen Diskurs – vor allem in den 1960er Jahren intensiv diskutiert, jedoch bis in die aktuellen Debatten hinein wirksam – unterschiedliche Positionen über die Zielsetzungen religiöser Bildung an Schulen auszumachen, die sich zwischen einer Bibelorientierung und einer Gesellschaftsorientierung bewegen: Während der Religionsunterricht vor 1960 in erster Linie an der Bibel orientiert war, wurde im sogenannten Problemorientierten Religionsunterricht eine konsequente Ausrichtung religiöser Bildung auf gesellschaftsrelevante Fragestellungen gefordert, was insbesondere auch die Urteilsfähigkeit und Mündigkeit der Schülerinnen und Schüler als ein wesentliches Ziel des Religionsunterrichts in den Vordergrund rückt.⁷ In der aktuellen Debatte über religi-

⁶ Ebd.

⁷ Vgl. bspw. HORST GLOY, Themen statt Texte? In: NORBERT SCHNEIDER (Hrsg.), Religionsunterricht. Konflikte und Konzepte. Beiträge zu einer neuen Praxis. Hamburg 1971, 67–79, insbesondere 74. Sowie der programmatische Titel »Muss die Bibel im Mittelpunkt des Religionsunterrichts stehen?« von HANS BERNHARD KAUFMANN zu finden in: FOLKERT RI-

onspädagogische Konzeptionen wird eine Synthese angestrebt, die sowohl die biblischen Traditionen als auch gesellschaftsrelevante Fragestellungen didaktisch berücksichtigt. Dies verdeutlichen PETER BIEHL und MARTIN ROTHGANGEL auch sprachlich, indem sie von didaktischen Strukturen anstatt von »starren« Konzeptionen sprechen.⁸

Das beides miteinander zu verknüpfen ist, wird auch durch die Abituraufgabenstellungen in keinem der untersuchten Bundesländer unterlaufen – überall lassen sich sowohl traditionsbezogene und biblische als auch gesellschaftsorientierte Aspekte in den Fragestellungen der Urteilsbildung finden. Dennoch sind klare Schwerpunktsetzungen erkennbar, welche – wie oben bereits als These formuliert – auch auf das Verständnis von Religionspädagogik seitens der Verfasserinnen und Verfasser der Abituraufgabenstellungen zurückzuführen sein könnten. So ist beispielsweise die vor allem in Ostdeutschland nach 1945 – durch die Verdrängung des Religionsunterrichts aus den Schulen – präsente Christenlehre als ein möglicher Sozialisationsfaktor ostdeutscher Lehrerinnen und Lehrer zu nennen, welcher sich an »Wort und Sakrament« ausrichtete und somit einen stärkeren Bekenntnisbezug bedingen kann.⁹ Die gesellschaftskritischen Perspektiven im Religionsunterricht zu betonen ist hingegen eine Forderung, welche vor allem im Rahmen der sogenannten 68er-Bewegung in Westdeutschland entstanden ist.¹⁰ MANFRED L. PIRNER konstatiert, dass ein genereller Zusammenhang zwischen »persönlichen Weltbildern, Lebensorientierungen und Wertvorstellungen von Lehrerinnen und Lehrern« und ihrem »professionelle[n] Denken und Handeln« in der Erziehungswissenschaft weitgehend unbestritten ist.¹¹ NIPKOW spricht von »nicht ausschmelzbaren weltanschaulich-religiösen Denkvoraussetzungen« auf denen die Pädagogik beruht.¹² PIRNER merkt jedoch auch kritisch an, dass sich für den deutschsprachigen Raum »ein eklatantes Forschungsdefizit im Hinblick auf die weltanschaulich-religiöse Orientierung

CKERS/BERNHARD DRESSLER (Hrsg.), Thematisch-problemorientierter Religionsunterricht. Aufbruch – Bewährung in der Praxis – Impulse. Neukirchen-Vluyn 2003, 145–151.

⁸ PETER BIEHL/MARTIN ROTHGANGEL, Konzeptionen und Strukturen, in: MICHAEL WERMKE/GOTTFRIED ADAM/MARTIN ROTHGANGEL (Hrsg.), Religion in der Sekundarstufe II. Ein Kompendium. Göttingen 2006, 183–218.

⁹ Vgl. MARTIN STEINHÄUSER, Art. Christenlehre, in: Wissenschaftliches Religionspädagogisches Lexikon im Internet (www.wirelex.de) 2016.

¹⁰ Vgl. MATTHIAS HAHN, Problemorientierter Religionsunterricht, in: Wissenschaftliches Religionspädagogisches Lexikon im Internet (www.wirelex.de) 2016.

¹¹ Vgl. MANFRED L. PIRNER, Religiosität und Lehrerprofessionalität. Ein Literaturbericht zu einem vernachlässigten Forschungsfeld, in: Zeitschrift für Pädagogik 59 (2013), 201–218, hier 201.

¹² KARL E. NIPKOW, Bildung in einer Pluralen Welt. Bd. II: Religionspädagogik im Pluralismus. Gütersloh 1998, 108.

von Lehrkräften sowie mögliche Zusammenhänge mit ihrer Professionalität konstatieren« lässt.¹³ So kann auch in Anknüpfung an die Ergebnisse der vorliegenden Studie der Einfluss der religiösen Sozialisierung und Überzeugungen seitens der Lehrkräfte auf ihre Einstellung zum Religionsunterricht sowie auf ihre Zielesetzungen von Urteilsbildungskompetenzen – gerade vor dem Hintergrund der oben skizzierten historischen Umstände – als ein weiterführendes Forschungsinteresse benannt werden.

Von den Erklärungsmodellen unbenommen ist, dass die Diversität an Fragestellungen ebenso wie länderspezifische Akzentsetzungen eine positive Vielfalt erzeugen, welche keinen negativen Einfluss auf die Vergleichbarkeit der Leistungen nimmt und daher – vor dem Hintergrund der bildungsadministrativen Standardisierungsbemühungen – zu schützen ist. Diese Vielfalt wird auch in der theologischen Ethik eingelöst, welche sich gerade dadurch auszeichnet, dass sie nicht auf spezifische Themen reduziert ist, sondern sich sowohl »allgemeinethischen« als auch religiösen Fragestellungen widmet (vgl. Kapitel 2.1).¹⁴

Neben diesen Differenzen, die als Ausdruck der Freiheit der Bundesländer für eine kontextsensible Ausgestaltung von Unterricht und Prüfung verstanden werden können, sind in der Analyse der Abituraufgabenstellungen Differenzen sichtbar geworden, die sich negativ auf die Vergleichbarkeit auswirken. Um diese Differenzen soll es im Folgenden gehen.

9.1.2 Differenzen als Grenzen für eine gerechte Leistungsbewertung

Die EPA wurden zwar von der Kultusministerkonferenz extern in einem »top-down-Prozess« implementiert, die Aufsicht über die Umsetzung bleibt jedoch den Ländern in ihren unterschiedlichen Prüfungskulturen überlassen.¹⁵ Ein Befund der vorliegenden Studie, der sich nicht unmittelbar auf Urteilsfähigkeit bezieht, sich jedoch mit anderen Befunden der *Educational Governance Forschung* deckt, besteht in der Feststellung, dass die bildungsadministrativen Vorgaben in Form der EPA in unterschiedlichem Maße Umsetzung finden und dass es mitunter auch zu Implementationsbrüchen kommt, die die Vergleichbarkeit der Abschlüsse zwangsläufig in Frage stellen.¹⁶ Während es also – wie im voran-

¹³ PIRNER, Religiosität und Lehrerprofessionalität, 213.

¹⁴ Vgl. ANDERSEN, Einführung in die Ethik, 14 f.

¹⁵ Vgl. KÜHN, Steuerung und Innovation, 309.

¹⁶ Vgl. KÜHN, Steuerung und Innovation; CORNELIA GRÄSEL/ILKA PARCHMANN, Implementationsforschung – oder: der steinige Weg, Unterricht zu verändern, in: Unterrichtswissenschaft 32 (2004), 196–214; sowie zusammenfassend KÜHN, Aufgaben in (zentralen) Abschlussprüfungen, 73–89. So kann man beispielsweise auch für das Fach Deutsch starke länderspezifischen Differenzen in der Ausgestaltung der Erwartungshorizonte in zentralen

gegangenen Kapitel herausgestellt – gilt, die Freiheit der Bundesländer zu wahren, durch verschiedene Aufgabenkulturen auf ihre spezifischen religiösen Kontexte eingehen zu können, sollten Differenzen, die unterschiedliche Anforderungsniveaus und Intransparenz erzeugen, minimiert bzw. behoben werden. Dabei geraten vor allem die Erwartungshorizonte in den Blick. Diese sollen eine doppelte Funktion erfüllen: Zum einen dienen sie dazu, die Auswertungsobjektivität zu erhöhen und somit die geforderte Transparenz und Vergleichbarkeit innerhalb eines Bundeslandes zu sichern. Zum anderen müssen die Rahmenbedingungen, aber auch das Anforderungsniveau der Erwartungshorizonte mit den Erwartungen anderer Bundesländer vergleichbar sein, um Gerechtigkeit in Bezug auf die Allokation von Berufsausbildungen und Studienplätzen zu gewährleisten.

Als ein Faktor, der die Transparenz der Abituranforderungen und Bewertungen wesentlich einschränkt, wurde in der Studie die inkonsequente Verwendung der Operatoren identifiziert (vgl. insbesondere Kapitel 7.5.2). Da die Operatoren in den EPA eindeutig definiert sind und ihre Verwendung geregelt ist, kann in Bezug auf diesen Faktor von einem »Implementationsbruch« gesprochen werden. Unstimmigkeiten in der Verwendung der Operatoren sind im Kontext von Abschlussprüfungen auch für das Fach Geschichte empirisch nachgewiesen.¹⁷ In den aus Thüringen stammenden Aufgabenstellungen und Erwartungshorizonten der hiesigen Materialien war die inkonsistente Nutzung der Operatoren besonders auffällig: Die wiederholt stattfindende Strukturierung der Erwartungshorizonte anhand von Operatoren, die in den Aufgabenstellungen keine Erwähnung finden und die wiederholt freie Auslegung der Operatoren durch die Erwartungshorizonte führen zu Intransparenz für die Prüflinge. So fordern die Erwartungshorizonte mitunter Handlungsschritte, die aus den Aufgabenstellungen nicht hervorgehen (vgl. Kapitel 7.5). Darüber hinaus sind den Prüflingen die Punkteverteilung und damit die Schwerpunktsetzungen der Erwartungshorizonte nicht ersichtlich, wodurch die aus Thüringen stammenden Aufgabensets unnötige Hürden für die Erfüllung der Erwartungshorizonte schaffen.

HERBERT ALTRICHTER und KATHARINA MAAG MERKI weisen darauf hin, dass es sich bei Steuerungsprozessen – wie sie die Umsetzung verbindlicher Prüfungsvorgaben darstellt – um multifaktorielle Prozesse auf mehreren Ebenen mit

Abschlussprüfungen für den Mittleren Schulabschluss empirisch nachweisen (vgl. STEFANIE MATHES/SVENJA M. KÜHN, »... dann kommt alles auf die Korrekturanleitung an...« Eine empirische Annäherung an ein [bislang] weitgehend vernachlässigtes Forschungsfeld, in: STEFAN KELLER/CHRISTIAN REINTJES [Hrsg.], Aufgaben als Schlüssel zur Kompetenz. Didaktische Herausforderungen, wissenschaftliche Zugänge und empirische Befunde. Münster 2016, 169–178).

¹⁷ Vgl. SCHÖNEMANN/THÜNEMANN/ZÜLSDORF-KERSTING, Was können Abiturienten?, 32.

unterschiedlichen Akteuren handelt, in denen die Vorgaben rekontextualisiert und adaptiert werden müssen.¹⁸ Erklärungsmodelle für »Brüche« müssen folglich den Kontext der Bundesländer jeweils mitbedenken: Möglicherweise ist die Inkonsistenz zwischen Aufgabenstellungen und Erwartungshorizonten der aus Thüringen stammenden Fälle durch die in diesem Bundesland geringe Relevanz des Faches für die Abiturprüfung zu erklären. Wie in Kapitel 5.1 dargestellt, legten in Thüringen im Jahr 2019 nur 17 Schülerinnen und Schüler eine schriftliche Abiturprüfung ab, während es in Bayern 339 und in Niedersachsen insgesamt 938 Schülerinnen und Schüler waren (gA: 389, eA: 549).¹⁹ Da es in ganz Thüringen nur an einer Schule möglich ist, eine schriftliche Prüfung im Fach Evangelische Religionslehre abzulegen, ist auch der Kreis derer geringer, die mit den Erwartungshorizonten arbeiten: Eine der beiden oben genannten zentralen Funktionen von Erwartungshorizonten, nämlich der, dass Auswertungsobjektivität innerhalb des Bundeslandes hergestellt werden muss, bleibt eine abstrakte Forderung, die zwar angesichts einer gerechten und transparenten Bewertung anzustreben ist, jedoch keiner realen Überprüfung unterzogen werden kann. Es ist somit denkbar, dass sich die Quantität der Prüflinge auf die Qualität der Erwartungshorizonte – gerade in Bezug auf die Umsetzung der Vorgaben – auswirken, auch wenn dies aus pädagogischer Sicht nicht haltbar ist, da Transparenz und Vergleichbarkeit auch für wenige Schülerinnen und Schüler angestrebt werden muss. Gleiches gilt für die korrigierenden Lehrerinnen und Lehrer, denen die Erwartungshorizonte eine gute Bewertungsbasis sein sollen, um eine möglichst objektive Auswertung vorzunehmen: Auch für sie ist die Passung zwischen Aufgabenstellungen und Erwartungshorizonten wichtig. Interessant wäre die Analyse der Erwartungshorizonte anderer Fächer, um zu prüfen, ob die Qualitätsunterschiede genereller oder fachspezifischer Art sind.

Eine weitere sichtbar gewordene Differenz, die es kritisch zu hinterfragen gilt, wird in Bezug auf den Aufgabenumfang im Verhältnis zur Bearbeitungszeit deutlich: Während die Prüflinge in Bayern 210 Minuten für durchschnittlich 6,92 Aufgaben haben, dürfen sich niedersächsische Prüflinge für durchschnittlich 4,13 Aufgaben 300 Minuten Zeit nehmen. Thüringen stellt mit durchschnittlich 2,75 Aufgaben bei 270 Minuten Bearbeitungszeit die meiste Arbeitszeit pro Aufgabe zur Verfügung.²⁰ In den vertiefenden Fallanalysen konnte

¹⁸ Vgl. HERBERT ALTRICHTER, KATHARINA MAAG MERKI, Steuerung der Entwicklung des Schulwesens, in: DIES. (Hrsg.), Handbuch. Neue Steuerung im Schulwesen. Wiesbaden ²2016, 1–27.

¹⁹ Dass die Zahlen nicht in erster Linie, aber auch durch die Unterschiede der Bevölkerungsmenge beeinflusst sind, wurde bereits in Kapitel 5.1 diskutiert.

²⁰ Die Gesamtzahl der zu bearbeitenden Aufgabenstellungen kann mit Ergebnissen aus den naturwissenschaftlichen Fächern kontrastiert werden: Die Studie von KÜHN kommt zu dem Ergebnis, dass in den Fächern Biologie, Chemie und Physik durchschnittlich 33,36 Opera-

gezeigt werden, dass dies – blickt man auf komplexe Urteilsbildungsaufgaben – mitunter zu erheblicher Zeitknappheit für bayerische Prüflinge führen kann, welche zu Ungerechtigkeiten führen, da sie die Vergleichbarkeit beeinträchtigen (vgl. Kapitel 7.2).

In Bezug auf diesen Faktor stellen jedoch die Vorgaben der EPA selbst die Weichen für diese Differenzen: Sie formulieren, dass die Anzahl der Teilaufgaben aufgrund der unterschiedlichen Bearbeitungszeiten differieren kann, ohne einen Richtwert mit einer Zeit zu verbinden. Vergleicht man die Bearbeitungszeiten, wird deutlich, dass Bayern trotz umfangreichster Aufgabenanzahl den Prüflingen die geringste Bearbeitungszeit gibt. Als Richtlinie für das erhöhte Anforderungsniveau werden in den EPA »bis zu sechs« Teilaufgaben genannt, so dass also die Bundesländer Thüringen und Niedersachsen trotz höherer Bearbeitungszeit im Toleranzrahmen der Vorgaben liegen, den Bayern mit durchschnittlich 6,92 Aufgaben geringfügig übersteigt.²¹

Diese scheinbar geringfügigen Abweichungen nach oben und die Offenheit der Formulierung »bis zu sechs« schaffen ungleiche Verhältnisse, welche mit Blick auf bayerische Abituraufgabensets dadurch verstärkt werden, dass sie im Gegensatz zu niedersächsischen und thüringischen Aufgabensets verschiedene Themenbereiche vereinen, so dass die Prüflinge ein breiteres Wissen nachweisen müssen und – obwohl ihnen vier statt nur zwei Wahlaufgabensets zur Verfügung stehen – unliebsame Themen schwerer umgehen können. Hinzukommt, dass das Sichten und das Auswählen der zur Verfügung stehenden Wahlaufgaben durch die Prüflinge zusätzlich Zeit in Anspruch nehmen.

Deutlich wird, dass die bundesweite Vergleichbarkeit, die man sich von der Einführung einheitlicher Prüfungsanforderungen und dem Zentralabitur innerhalb der Bundesländer erhoffte, hinsichtlich des abgefragten Umfangs und der Bearbeitungszeit nicht gegeben ist.

Anhand der aus Thüringen stammenden Aufgabensets ist jedoch auch zu beobachten, dass eine geringere Zahl an Aufgabenstellungen nicht zwangsläufig schwierigkeitsreduzierend wirken muss, da mit ihr in der Regel auch ein weniger kleinschrittiges Vorgehen verbunden ist, welches zu einer höheren Komplexität der einzelnen Aufgaben führen kann. So lassen sich beispielsweise in Niedersachsen alle Teilaufgaben unter ein übergeordnetes Thema stellen, so dass die Prüflinge durch die vorangegangenen Aufgabenstellungen in der Urteilsbildung

torenaufgaben von den Prüflingen gelöst werden müssen (vgl. KÜHN, Steuerung und Innovation, 301).

²¹ Für Abiturprüfungen mit grundlegendem Anforderungsniveau sind bis zu vier Teilaufgaben als Richtwert genannt.

Zur Funktion der Teilaufgaben benennen die EPA zum einen die Eingrenzung, Akzentuierung und Präzisierung der Prüfungsaufgabe sowie die Eröffnung verschiedener Blickwinkel, die Förderung von Vernetzung und die Forderung verschiedener AFB (vgl. EPA, 18).

entlastet werden, da sie inhaltliche Aspekte vorsondieren, ordnen oder vergleichen müssen. Solch ein kumulativer Aufbau der Aufgabensets fehlt in Thüringen in mehreren Fällen, so dass die Aufgaben sehr offen und komplex werden, wie in der vertiefenden Einzelfallinterpretation exemplarisch gezeigt werden konnte (vgl. Kap. 7.5.1).

9.1.3 Die Kunst der Unterscheidung

Die Unterscheidung zwischen einer gewünschten kontextsensiblen Vielfalt und einheitlichen Standards wirft die Frage nach der Grenzziehung auf: Ab wann ist eine Standardisierung notwendig und wo handelt es sich um angemessene Unterschiede? Im Datenkorpus der vorliegenden Studie lassen Urteilsbildungen im Zusammenhang mit dem Apostolischen Glaubensbekenntnis in einer Aufgabe in Thüringen und einer Aufgabe in Bayern einen interessanten Vergleich zu, der das Problem illustriert: In Thüringen, wird den Schülerinnen und Schülern das Bekenntnis als Materialgrundlage vorgelegt, während in Bayern davon ausgegangen wird, dass die Prüflinge das Apostolikum (auswendig) kennen [TH, 2018, eA (1) und BY, 2019 (4)]. Die Anforderungen sind in Bayern folglich höher als die in Thüringen. Auch hier könnte der religionssoziologische Kontext zu den Annahmen bei den Verfasserinnen und Verfassern der Aufgabenstellungen führen, dass das Apostolische Glaubensbekenntnis unter den Schülerinnen und Schülern in Bayern vorausgesetzt werden kann, während die Verfasserinnen und Verfasser thüringischer Abituraufgabenstellungen nicht davon ausgehen, dass das Apostolische Glaubensbekenntnis im konkreten Wortlaut – den eine gründliche Analyse voraussetzt – hinreichend bekannt ist. Blickt man jedoch auf das einzelne Individuum, welches in Bayern ebenso säkular erzogen worden sein kann wie in Thüringen, erscheint der Unterschied ungerechtfertigt.²²

Die Ungleichheiten, die durch Materialienbeigaben entstehen, sind jedoch schwer zu beheben. Es wäre maximal ein Kanon an Texten denkbar, welcher bei Prüfungen vorauszusetzen ist, wobei eine solche Engführung auf einzelne Texte angesichts der Breite an verfügbarem Wissen für viele Kritiker als eine ungerechtfertigte Tradierung einzelner »klassischer Texte« erscheint (vgl. Kapitel 3.1).

Hinsichtlich der aufgezeigten Differenzen muss das leitende Kriterium sein, ob sich diese negativ auf die Vergleichbarkeit der Abschlüsse auswirken, oder nicht. Diesbezüglich sind eindeutige formale Richtlinien ein wesentlicher Punkt der Qualitätssicherung, doch auch inhaltliche Kompetenzbeschreibungen

²² Da das Apostolikum in den Oberstufenlehrplänen der beiden Länder gleichermaßen unerwähnt bleibt, lassen sich auch wenig Rückschlüsse auf den vorangegangenen Unterricht ziehen, der den Autorinnen und Autoren der Abiturprüfungen jedoch auch nicht bekannt ist.

innerhalb der EPA können hilfreich sein, wie im folgenden Kapitel gezeigt werden soll.

9.2 Aspekte der Zielvorgaben – Gütekriterien und Bewertungsmaßstäbe für Urteilsfähigkeit in den Erwartungshorizonten

Neben den Vorgaben in Bezug auf Gestaltungsaspekte machen die EPA Vorgaben zum Outcome in Form von fünf Kompetenzen, welche in den Abiturprüfungen im Fach Religion geprüft werden sollen. Diese Vorgaben beziehen sich auf Teilaspekte der Kompetenz (vgl. Kapitel 3.2). Kompetenzspezifische Qualitätsmerkmale oder Niveaustufendifferenzierungen werden von den EPA nicht vorgegeben, so dass das zu messende Konstrukt, beispielsweise von Urteilsbildung, unscharf bleibt. Die Erwartungshorizonte füllen diese Lücke nur teilweise: Es werden zwar einige Qualitätsmerkmale genannt – eine für die Lernstandsdiagnostik notwendige hierarchische Strukturierung innerhalb der genannten Merkmalsausprägungen ist in den Erwartungshorizonten jedoch nicht gegeben (zur Diagnostik vgl. Kapitel 3.3). Der Konsens zwischen den drei untersuchten Bundesländern über die Qualitätsmerkmale, anhand derer eine Urteilsbildung zu bewerten sein sollte, ist zudem als gering einzuschätzen: Die in mindestens zwei der Bundesländer genannten allgemeinen Qualitätsmerkmale für Urteilsbildung waren neben dem Vorhandensein einer Stellungnahme, einer Begründung und dem Aufgabenbezug die fünf Merkmale »Differenziertheit«, »Eigenständigkeit« »fundierte Kenntnisse/sachgerecht«, »Schlüssigkeit« und »Verständlichkeit/sprachliche Gestaltung«.²³ Die genannten Qualitätsmerkmale werden im Folgenden aus einer diagnostischen Perspektive betrachtet (Kapitel 9.2.1), bevor sie in den Kontext weiterer Merkmale gestellt werden, die im Forschungsdiskurs über Urteilsfähigkeit diskutiert werden (vgl. Kapitel 9.2.2). Kapitel 9.2.3 diskutiert abschließend die Ergebnisse der Studie vor dem Hintergrund messtheoretischer Ansprüche.

²³ In der Studie wurde »allgemeine Qualitätsmerkmale« von »aufgabentextspezifischen Qualitätsmerkmalen« abgegrenzt, wobei sich letztere auf explizit in der Aufgabenformulierung genannte Merkmale beziehen (vgl. Kap. 6.3.2).

9.2.1 Qualitätsmerkmale der Erwartungshorizonte – Differenziertheit, Eigenständigkeit, fundierte Sachkenntnisse, Schlüssigkeit und sprachliche Gestaltung

Die abstrakten und im Allgemeinen verbleibenden Formulierungen der Qualitätsmerkmale – die in den vertiefenden Fallanalysen wiederholt aufgezeigt wurden – sollen hier noch einmal an konkreten Merkmalen problematisiert und diskutiert werden. So ist beispielsweise für das Kriterium »Differenziertheit«, welches mit 16 Nennungen in den insgesamt 42 untersuchten Erwartungshorizonten eines der am häufigsten genannten Qualitätsmerkmale darstellt, wiederholt der Bezugspunkt unklar: Differenziertheit kann sich auf verschiedene Teilkompetenzen von Urteilskompetenz beziehen, wie beispielsweise die Folgenreflexion oder auf das Erfassen und Reflektieren verschiedener Teilaspekte des ethischen Problems. Ebenso kann der Rückbezug zu ethischen Konzeptionen, Normen und Werten, die in der Fragestellung verankert sind, »differenziert« sein (zu den Teilkompetenzen von Urteilsfähigkeit vgl. Kapitel 4, insbesondere 4.1.4).

Das Kriterium ist darüber hinaus – wie bereits anhand der vertiefenden Einzelfallanalyse aus Bayern gezeigt werden konnte – von der Komplexität der Fragestellung abhängig: Je komplexer der Sachverhalt, umso differenzierter kann die Begründung einer Urteilsbildung ausfallen (vgl. Kapitel 7.3.1). Das Kriterium hängt folglich vom (fachlichen) Durchdringungsgrad ab, welcher an die Komplexität und Abstraktion der Fragestellung gekoppelt ist. In mehreren Erwartungshorizonten zu Urteilsbildungsaufgaben fehlen jedoch inhaltlich hinreichend determinierte Verstehensziele, durch welche man sich auf ein Kriterium einigt, wann beispielsweise die ethische Dimension oder eine Folgenabschätzung vom Prüfling als »differenziert erfasst« gelten kann (vgl. die Fallinterpretationen in Kapitel 7). Da die inhaltlichen Aspekte zur gestellten Fragestellung in den Erwartungshorizonten oft unzureichend bestimmt sind und es auch kaum Differenzierungen zwischen inhaltlichen Aspekten gibt, die für die Lösung der Aufgabe zentral sind, und jenen, die optionalen oder exemplarischen Charakter haben, obliegt es der fachlichen Kompetenz der korrigierenden Lehrkräfte zu entscheiden, was als eine »differenzierte Betrachtung« gilt und was nicht bzw. auf welcher Niveaustufe sich die Differenziertheit befindet. Die Verbindung zwischen unzureichenden inhaltlichen Merkmalen und der Allgemeinheit des Kriteriums »Differenziertheit« macht es wahrscheinlicher, dass die Expertise der Lehrkraft in Bezug auf die Themenstellung zum Maßstab wird, was zu Messfehlern führen kann: Aus der pädagogischen Diagnostikforschung ist beispielsweise der Kontrast- bzw. Ähnlichkeitsfehler bekannt. Er beschreibt einen Messfehler, bei dem sich die persönliche Begabung der Lehrkraft auf das Messergebnis auswirkt: Lehrkräfte mit einer besonderen Expertise auf einem Gebiet bewerten die Ergebnisse demnach schlechter als durchschnittlich infor-

mierte Lehrkräfte.²⁴ Wenn das Kriterium »Differenziertheit« keine inhaltlichen Bezugspunkte hat, liegen solche Fehler nahe. Für den Kontrast- bzw. Ähnlichkeitsfehler ist auch das Kriterium »fundierte Sachkenntnisse« anfällig, sofern dieses nicht im Erwartungshorizont expliziert wird.

Ähnlich schwer zu greifen ist das Qualitätsmerkmal »Eigenständigkeit«, welches in 15 der 42 analysierten Erwartungshorizonte genannt wird und hauptsächlich in zwei verschiedenen Verbindungen vorkommt: Zum einen ist eine »eigenständige Stellungnahme oder Urteilsfähigkeit«, zum anderen eine »eigenständige Begründung« gefordert. Was bedeutet es, wenn ein Urteil als »eigenständig« angesehen werden kann? Grundsätzlich ist Eigenständigkeit das Gegenteil einer unreflektierten Übernahme. Vor allem im Kontext des bekenntnisorientierten Religionsunterrichts kann es als ein Signal dahingehend gelesen werden, dass eine christliche oder anderweitig vorgegebene Position nicht übernommen werden muss, sondern der Prüfling in seinem Urteilsentscheid frei und selbstständig ist. Messtheoretisch ist dieses Kriterium schwer zu fassen. DIRK FELZMANN weist darauf hin, dass es einen Unterschied macht, ob »fachwissenschaftliche[] Urteile aus dem Unterricht bekannt sind und damit nur reproduziert werden müssen oder ob sie eigenständig auf Basis der bereitgestellten Informationen und Unterrichtswissen neu konstruiert werden müssen«²⁵. Im Kontext zentral gestellter Abschlussprüfungen kann diese Differenzierung jedoch nicht erfasst werden. Da der Mehrheit der Aufgabenstellungen Material zugrunde liegt, welches auch häufig Positionierungen zur Fragestellung enthält – die EPA fordern sogar, dass aus den zugrunde liegenden Materialien eine Positionierung hervorgeht²⁶ – tritt unter diesem Qualitätsmerkmal der Umgang mit den anderen Positionen in den Fokus: Eigenständigkeit erfordert, sich anderer Positionen bewusst zu sein, sich jedoch auch von diesen abgrenzen zu können und sie abzuwägen.

Das Kriterium »Schlüssigkeit«, welches in sechs der 42 untersuchten Erwartungshorizonte genannt wird, spricht die Grundzüge formaler Logik an. Aus argumentationstheoretischer Perspektive entsteht Schlüssigkeit, wenn die semantische Verbindung zwischen These und Argument überzeugend ist (vgl. zu den Grundlagen der Argumentationstheorie Kapitel 4.1.1). Es kommt folglich nicht nur auf das Vorhandensein von Argumenten, sondern auch auf ihre Qualität und Situationsangemessenheit an. IRIS WINKLER problematisiert das Phänomen zu allgemeiner Qualitätskriterien unter anderem am Beispiel der Schlüssigkeit für die Bewertung von Argumentationen und macht durch ihre Aussage, die sich auf die Deutschdidaktik bezieht, gleichzeitig deutlich, dass das Formulieren von

²⁴ Vgl. VON SALDERN, Schulleistung 2.0, 101.

²⁵ DIRK FELZMANN, Das Bewerten bewerten. Klausuraufgaben zur Beurteilungs- und Bewertungskompetenz, in: Geografie und Schule, 35 (2013) 206, 15–22, hier 18.

²⁶ EPA, 16.

Erwartungshorizonten für komplexe Aufgabenstellungen eine fächerübergreifende Herausforderung darstellt:

»Der bloße Hinweis, dass Argumentationen »schlüssig«, Argumente »stichhaltig« oder Beispiele »treffend« sein müssten, hilft in diesem Zusammenhang nicht weiter. Wenn nicht explizit gemacht wird, welche Anforderungen ein Argument erfüllen muss, um plausibel zu sein, hängt die Betonung dieses Aspekts zu stark von der Lehrperson ab, die u. U. ihrerseits nur den formal-strukturellen Blick auf Argumentationen gelernt hat.«²⁷

Statt des »formal-strukturellen« Blicks betont WINKLER, dass Argumente an die Adressaten und an den Kontext angepasst sein müssen, um wirksam zu sein.²⁸ Ihre Aussage ist jedoch auch als ein Plädoyer dafür zu lesen, die Anforderungen an Argumente deutlich zu machen. Dies bedarf eines verstärkten Diskurses über Qualitätsmerkmale und einer ausgeprägten Sprachfähigkeit über Anforderungen, die sowohl im Lehramtsstudium, im Referendariat und auch unter Lehrkräften forciert werden muss. In diesen Diskurs wären auch Qualitätsmerkmale aufzunehmen, die in den Erwartungshorizonten derzeit noch nicht oder nur randständig erscheinen, in der Forschungsliteratur jedoch über verschiedene Disziplinen hinweg als zentral angesehen werden. Sie werden im folgenden Kapitel betrachtet.

9.2.2 Gütekriterien und Bewertungsmaßstäbe in der Forschungsliteratur – Multiperspektivität, Transparenz der zugrunde liegenden Werte, Kontexteinbettung, Adressatenbezug, Folgenreflexion

Betrachtet man die in den Erwartungshorizonten des Untersuchungszeitraums genannten Qualitätsmerkmale vor dem Hintergrund der in der Forschungsliteratur diskutierten Merkmale (vgl. Kapitel 4), so sind einige Akzentverschiebungen sichtbar. In Bezug auf folgende vier Qualitätskriterien scheinen die Forderungen der Forschungsliteratur einerseits und die Erwartungshorizonte andererseits besonders zu divergieren: Es handelt sich um »Multiperspektivität«, »Transparenz der zugrunde liegenden Werte«, »Kontexteinbettung und Adres-

²⁷ Vgl. IRIS WINKLER, Argumentierendes Schreiben im Deutschunterricht im Spiegel von Aufgaben für Lern- und Leistungssituationen, in: ELKE GRUNDLER/RÜDIGER VOGT (Hrsg.), Argumentieren in Schule und Hochschule. Interdisziplinäre Studien. Tübingen 2006, 157–166, hier 160 f.

²⁸ Ebd.

satenbezug« sowie die »Folgenreflexion«. Diese vier Merkmale werden im Folgenden diskutiert.

Multiperspektivität

Während in der Forschungsliteratur das Prüfen von Gegenargumenten als ein wesentliches Qualitätsmerkmal benannt wird (beispielsweise GEORG LIND, ELKE VISSER, THERESA SCHWARZKOPF etc. – siehe Kapitel 4), taucht dieses in den Erwartungshorizonten nur selten explizit auf: In den aus Bayern stammenden Erwartungshorizonten wird es zweimal generell genannt und einmal vor dem Hintergrund einer expliziten Forderung in der Aufgabenstellung. In den aus Thüringen stammenden Erwartungshorizonten wird es insgesamt einmal genannt. In den Erwartungshorizonten aus Niedersachsen taucht es nicht explizit auf, jedoch wird in zwei Fällen eine »abwägende Prüfung« verlangt, was das Prüfen von Gegenargumenten voraussetzt.

Man könnte Multiperspektivität über diese wenigen Nennungen hinaus hinter dem Merkmal »Differenziertheit« vermuten, welches häufig in der Formel »überzeugende und differenzierte Begründung« zu finden ist, oder auch in dem Merkmal »Durchdringungstiefe«, welches in den Erwartungshorizonten aus Niedersachsen achtmal vorkommt. Jedoch hängt es an der Interpretation der Korrektorinnen und Korrektoren, diese Merkmale hinsichtlich der Prüfung von Gegenargumenten zu deuten. Es ist möglich, dass ein stilles Einvernehmen darüber besteht, dass dialektische Begründungen linearen vorzuziehen sind und dass es eine Selbstverständlichkeit darstellt, dass bei der schriftlichen Begründung eines Urteilsentscheides auch Gegenargumente diskutiert und Grenzen des Urteilsentscheids aufgezeigt werden. Wahrscheinlicher ist jedoch, dass in den Urteilsbildungsaufgaben des schriftlichen Abiturs im Fach Evangelische Religionslehre die Multiperspektivität der Begründung keine tragende Rolle spielt, denn in den Fällen, in denen konträre Positionen in Hinblick auf die Fragestellung ausformuliert sind, wird nicht die Abwägung dieser gefordert, sondern stets nur die Positionierung zu einer der genannten Positionen (vgl. beispielsweise den Erwartungshorizont in Kapitel 7.4.1). Wie mit der Position der Gegenseite umzugehen ist, wird nicht thematisiert.

In den von den EPA genannten Bewertungskriterien fehlt das Prüfen von Gegenargumenten ebenso als ein Qualitätsmerkmal. Dort wird zwar auf die »Breite der Argumentation«, die »Vielfalt der Aspekte« sowie die »Reichhaltigkeit der Argumente« verwiesen (vgl. Anhang II).²⁹ Diese Merkmale können jedoch auch in einer linearen Argumentation Umsetzung finden. Multiperspektivität wird nicht explizit gefordert oder bedacht.

²⁹ Vgl. EPA, 20.

Transparenz der zugrunde liegenden Werte

Ein weiteres in der Forschungsliteratur hervorgehobenes Merkmal ist die Offenlegung der Werte und Normen, die dem Urteilsentscheid zugrunde liegen. Im Untersuchungszeitraum ist dieses Kriterium in den drei untersuchten Bundesländern nur einmal genannt [E-NI, 2014, gA (1)]. Dieser Erwartungshorizont geht explizit auf die Möglichkeit ein, der in der Textgrundlage gegebenen Position zuzustimmen oder sie abzulehnen.³⁰ In Bezug auf die Ablehnung der dargestellten Positionierung heißt es: Bei einem Widerspruch »[...] ist es erforderlich, das jeweilige ethische Grundprinzip zu verdeutlichen, von dem aus argumentiert wird, [...]« [E-NI, 2014, gA (1)]. In allen anderen Fällen sind vage Formulierungen wie »Reflektiertheit«, »Differenziertheit«, »Begründung« oder »Schlüssigkeit« genannt, die die Transparenz der Wertgrundlagen selbstverständlich begrüßen, jedoch nicht explizit einfordern. Auch inhaltlich kommen die Erwartungshorizonte nicht auf die Abstraktionsebene, die divergierenden Normen und Werte, die einer kontroversen Fragestellung zugrunde liegen, zu benennen und zu reflektieren. Aus fachwissenschaftlicher Sicht bleibt damit ein wesentliches Merkmal von Urteilsbildung unbeachtet.

SCHÖNEMANN, THÜNEMANN und ZÜLSDORF-KERSTING kommen in einer empirischen Untersuchung von Abituraufsätzen im Fach Geschichte zu dem Ergebnis, dass in ihrem Datenkorpus keine Schülerlösung zu finden ist, in welcher der auch in der Geschichtsdidaktik präsente Anspruch, die Maßstäbe einer Urteilsbildung offen zu legen und zu reflektieren, eingeholt wird.³¹ FELZMANN konstatiert für den Geographieunterricht ebenfalls, dass das Offenlegen eines eigenen Wertmaßstabs zwar in den Kerncurricula Niedersachsens für die Sekundarstufe II verlangt wird, dass es jedoch für Klausuraufgaben eher untypisch ist.³² Er weist darauf hin, dass das Benennen von Wertmaßstäben auf die Schülerinnen und Schüler konstruiert wirken kann, da gemeinsam geteilte Werte in der alltäglichen Kommunikation normalerweise nicht expliziert werden³³. In Bezug auf kontroverse Fragestellungen sind es jedoch die konfligierenden Werte und Normen, die die ethische Dimension der Fragestellung erst bewirken (vgl. Kapitel 2.1).³⁴ Das Verständnis der der Fragestellung zugrunde liegenden Wer-

³⁰ Auch hier wird die oben diskutierte Multiperspektivität nicht angesprochen. Eine Positionierung im Sinne von Zustimmung oder Ablehnung scheint aus der Logik des Erwartungshorizontes ausreichend, ohne dass die explizite Betrachtung der Gegenargumente notwendig ist.

³¹ Vgl. SCHÖNEMANN/ THÜNEMANN/ ZÜLSDORF-KERSTING, Was können Abiturienten?, 70.

³² Vgl. FELZMANN, Das Bewerten bewerten, 19.

³³ Vgl. a. a. O., 20.

³⁴ Vgl. dazu auch FELZMANN, der ebenso auf die Erarbeitung miteinander in Konflikt stehender Werte abhebt: ebd. Vgl. auch die Teilkompetenz »Wahrnehmen und Bewusstmachen der moralischen Relevanz« bei ALFS et al., Bewertungsprozesse, 87–91.

tekonflikte ist eine Voraussetzung dafür, diese Werte zu gewichten und sich selbst zu einem Wert zu positionieren – also seine eigenen Wertvorstellungen offen zu legen. Die Studie hat jedoch gezeigt, dass dies auf der Ebene der Erwartungshorizonte nicht erfolgt: Weder konfligierende Werte noch kontroverse Position finden in der Regel eine explizite Ausformulierung und Reflexion, was sowohl durch Kategorie H *Formulierung konträrer Positionen in den Erwartungshorizonten* (vgl. Kap. 6.3.4) als auch durch eine vertiefende Einzelfallanalyse (vgl. exemplarisch Kap. 7.3.3) herausgearbeitet werden konnte. Dadurch, dass die konträren Sichtweisen nicht benannt sind, fehlt vielen Lösungsskizzen eine Verständigung darüber, worin die Strittigkeit der aufgeworfenen ethischen oder religiösen Fragestellung liegt, die vom Prüfling herausgearbeitet und dann bearbeitet werden soll. Es wäre ein großer Gewinn für die Erwartungshorizonte, wenn sie die verletzten oder zur Diskussion stehenden Werte klar benennen und das Erarbeiten der ethischen Dimension des Problems als einen essentiellen Bestandteil der Lösung markieren würden, da – und damit ist ein weiteres generelles Problem angesprochen – in den aktuellen Erwartungshorizonten in der Regel unklar bleibt, welche der genannten Aspekte zur Lösung der Aufgabe essentiell sind und welche optionalen oder exemplarischen Charakter haben.³⁵

Kontexteinbettung und Adressatenbezug

Das in der Kompetenzorientierung geforderte Prinzip der lebensweltlichen Anforderungssituationen hat auch in den Abituraufgabenstellungen durch Gestaltungsaufgaben Einzug gehalten, so dass das Kriterium »Kontexteinbettung und Adressatenbezug« in den Aufgabenstellungen und Erwartungshorizonten wiederholt auftaucht. 17 der 42 Urteilsbildungsaufgaben sind in einen Kontext eingebettete Gestaltungsaufgaben (~ 40%). In den Erwartungshorizonten wird dieser Kontext durch allgemein gefasste Stichpunkte wie »Adressaten-«, »Situations-« oder »Materialbezug« aufgegriffen. Aus wissenschaftlicher Perspektive wird im Kontext von Urteilsfähigkeit hervorgehoben, dass bestimmte Argumente ihre Gültigkeit nur in bestimmten Situationen und nur in Bezug auf

³⁵ So kann die von THOMAS ZABKA und TOBIAS STARK für Korrekturanweisungen des Faches Deutsch erhobene Forderung, dass der »jeweilige aspektbezogene Grad der Verbindlichkeit genauer erläutert und für die Korrektoren transparent gemacht wird«, auch auf das Fach Religion übertragen werden (vgl. THOMAS ZABKA/TOBIA STARK, Aufgabenstellungen und Erwartungshorizonte als Steuerungsinstrumente. Zum Umgang mit Problemen der Literaturinterpretation im Zentralabitur, in: Der Deutschunterricht 1 (2010), 19–29, hier 28). Damit würde erreicht, dass eine Unterscheidung möglich wird, zwischen »(a) Erwartungen, die diesseits der Deutungsoffenheit liegen und inhaltlich konkret bestimmbar sind, und (b) Erwartungen, die sich auf einen offenen oder in bestimmten Grenzen offenen Entscheidungsspielraum richten.« (vgl. ebd.).

bestimmte Argumentationsteilnehmerinnen und -teilnehmer entfalten.³⁶ Gefordert wäre demnach eine Reflexion des Kontextes hinsichtlich möglicher Argumentationsstrategien, die auf *Leserbewusstheit*, *Lesereinschätzung* oder *Leseranpassung* (THOMAS JECHLE) schließen würden.³⁷ Das Merkmal Adressatenbezug ist folglich dahingehend zu hinterfragen, ob es hinsichtlich der Argumentationsmuster einen Unterschied machen sollte, ob ein Brief an den Theologen NIKOLAUS SCHNEIDER [NI, 2016, gA (1)] oder ein Artikel für die Schülerzeitung [BY, 2017 (3)] verfasst werden soll. Wenn das Kriterium »Adressatenbezug« ausschließlich durch die Einsetzung einer der Textform entsprechenden Anrede eingelöst ist, scheinen Gestaltungsaufgaben als »innovative Aufgabenformate« in den Abitursets einen Selbstzweck zu erfüllen. Die These, dass es sich bei der Forderung eines Adressatenbezugs lediglich um die Übernahme einer didaktisch postulierten Anforderung handelt, wird auch an einem Erwartungshorizont im Untersuchungskorpus deutlich, in welchem ein Adressatenbezug gefordert wird, ohne dass die Aufgabenstellung einen Adressaten benennt [E-TH, 2014 (2)] (vgl. Kapitel 7.5.1). Eine Metareflexion der vorgegebenen Kontexte fehlt in den Erwartungshorizonten, wobei empirisch zu prüfen wäre, welche Leistungen man von Abiturientinnen und Abiturienten, die auf diese Anforderungen im Religionsunterricht vorbereitet wurden, realistisch erwarten kann, um diese vor zu hohen Anforderungen zu schützen.

Folgenreflexion

Eine letzte zu diskutierende Diskrepanz zwischen den Forderungen der Forschung zu Urteilsfähigkeit und den Erwartungshorizonten der Abituraufgabenstellungen ist die Bedeutung, die einer Folgenreflexion beigemessen wird. ELKE VISSER definiert die Berücksichtigung gesamtgesellschaftlicher Folgen als die höchste von drei Niveaustufen dieses Kriteriums, während auf Niveaustufe II nur die Folgen auch mittelbar Betroffener und auf Niveaustufe I nur die Folgen des Urteils auf unmittelbar Betroffene berücksichtigt werden (vgl. Kapitel 4.1.4).³⁸

In den analysierten Erwartungshorizonten ist eine Folgenreflexion nur ein randständiges Thema, wobei das Kriterium für theologische und metaphysische Fragen auch eine geringere Relevanz besitzt als beispielsweise in Bezug auf bioethische Fragestellungen. Dennoch kann die von Visser vorgeschlagene graduelle Einstufung den Diskurs über Qualitätskriterien von Urteilsbildung und die Planung von Lernarrangements in Bezug auf ethisch-philosophische Ur-

³⁶ TOULMIN, Gebrauch von Argumenten, 37–39.

³⁷ JECHLE arbeitet diese Bezüge als Qualitätsmerkmale in einer Studie über kommunikatives Schreiben heraus (vgl. THOMAS JECHLE, Kommunikatives Schreiben. Prozeß und Entwicklung aus der Sicht kognitiver Schreibforschung. Tübingen 1992).

³⁸ VISSER, Diagnose der Bewertungskompetenz, 289.

teilsbildungen bereichern, da die Folgen bisher kaum und vor allem nicht in graduellen Abstufungen von den Erwartungshorizonten aufgegriffen werden.

9.2.3 Die Erwartungshorizonte vor dem Hintergrund von Messtheorien

Die in den Erwartungshorizonten genannten Qualitätsmerkmale und die geforderten Anforderungen machen eine fehlende Sprachfähigkeit in Bezug auf eine graduelle Beschreibung von Qualitätsmerkmalen für Urteilsfähigkeit deutlich. Die verwendeten Begriffe vermögen es nicht, einen messbaren Standard zu benennen, der in den Aufgabenstellungen geprüft und bewertet werden soll. Es mangelt zum einen an einem an den Forschungsstand angepassten Konzept von Urteilsfähigkeit (A). Zum anderen an einer Systematik der Leistungsbeschreibung, die eine Hierarchisierung der Leistungen zulässt (B). Solch ein Standard wäre für die Qualitätssicherung der Förderung von Urteilskompetenz in den EPA zu verankern und für spezifische Aufgabenstellungen in den Erwartungshorizonten auszuformulieren.

(A) Ein Strukturmodell zur Urteilsbildung in den EPA

Bildungsstandards dienen dazu, die »Einlösung der Anforderungen [zu] überprüfen. So lässt sich feststellen, inwieweit das Bildungssystem seinen Auftrag erfüllt hat, und die Schulen erhalten eine Rückmeldung über ihre Arbeit.«³⁹ Dies gilt auch für den einleitend beschriebenen Anspruch des Religionsunterrichts »die Auseinandersetzung mit Ideologien, Weltanschauungen und Religionen [zu] ermöglichen«, »für die Verständigung über gesellschaftliche Grundorientierungen unverzichtbar« zu sein und die Schülerinnen und Schüler dazu zu befähigen, »in religiösen und ethischen Fragen begründet urteilen« zu können.⁴⁰ Das schließt ein, dass der Religionsunterricht sprachfähig darüber ist, wie diese Ansprüche und Anforderungen einzuholen sind. Diese Benennung von Standards, die auf dem wissenschaftlichen Erkenntnisstand fußen und an diesen kontinuierlich angepasst werden müssen, ist zum einen für die Transparenz gegenüber den Schülerinnen und Schülern und zur Evaluierung des Ertrags unterrichtlicher Bemühungen von Bedeutung, aber zum anderen auch für die Religionslehrerbildung. Diese kann die Ausbildung kompetenzorientierter Lehrerinnen und Lehrer nur dann erfüllen, wenn sie die Fragestellung ins Zentrum rückt, woran ein Kompetenzerwerb der zentralen Kompetenzen Wahrnehmungs- und Darstellungsfähigkeit, Deutungsfähigkeit, Urteilsfähigkeit, Dialogfähigkeit und Gestaltungsfähigkeit in Schülerausagen und -leistungen sichtbar wird und wie Schülerinnen und Schüler »auf ihrem

³⁹ BMBF (Hrsg.), Bildungsstandards, 19.

⁴⁰ Vgl. EPA, 5 und 8.

jeweiligen Entwicklungsniveau [...] auf die nächste Zone ihrer Entwicklung [...]»⁴¹ gehoben werden können. Auf den Kerngegenstand der Studie bezogen, führt dies zu der Fragestellung: Wie schaffe ich als Lehrkraft Lernangebote, die meine Schülerinnen und Schüler zu einer differenzierten, eigenständigen, fundierten und schlüssigen Urteilsbildung befähigen? Zugrunde liegt die Kernfrage, was ein differenziertes, eigenständiges, fundiertes und schlüssiges Urteil in einer ethischen oder religiösen Fragestellung ist. Die Benennung wesentlicher Qualitätsmerkmale ethischer und religiöser Urteilsbildung in den EPA könnte den Diskurs über solche Standards anregen und die Sprachfähigkeit in Bezug auf die Formulierung von Lernzielen und Standards steigern. Vor dem Hintergrund der Einordnung der Analyseergebnisse in den Kontext der Forschungsliteratur sollten dabei folgende fünf, derzeit in den Erwartungshorizonten noch randständigen Aspekte als Anregungen zur Qualitätssicherung in die EPA aufgenommen werden:

1. Das Herausarbeiten der ethischen Dimension durch die Benennung verletzter/zur Diskussion stehender Werte
2. Multiperspektivität: neben der Benennung kontroverser Positionen auch die Wahrnehmung und Auseinandersetzung mit Gegenpositionen
3. Die Reflexion des zu berücksichtigenden Materials hinsichtlich seiner Aussagekraft
4. Die Folgenreflexion in Bezug auf ethische Fragen
5. Die explizite Betonung der Offenheit des Urteilsentscheides (vgl. dazu auch Kapitel 9.3).

Insgesamt ist aus messtheoretischer Sicht eine explizitere Ausformulierung der Anforderungen an Urteilsfähigkeit zu fordern, da diese in vielen analysierten Erwartungshorizonten nur allgemein und unzureichend formuliert sind.

(B) Eine Systematik der Hierarchisierung der Leistung – einen Weg zwischen synthetischer und analytischer Leistungsbeurteilung finden

Ein weiterer Orientierung-gebender und Auswertungsobjektivität-steigernder Faktor ist die Hierarchisierung der benannten Qualitätsmerkmale. Er macht eine Bewertung mittels der Vergabe von Bewertungseinheiten – wie sie im schriftlichen Abitur verlangt wird – erst möglich. Zum einen kann dabei auf bereits ausgearbeitete Niveaustufenformulierungen, wie sie beispielsweise von ELKE VISSER für die Biologiedidaktik empirisch überprüft wurden, zurückgegriffen werden (vgl. Kapitel 4.1.4). Diese Art von inhaltlich definierten Niveaustufenmodellen wären in den religionspädagogischen Diskurs aufzunehmen, um die oben geforderte Sprachfähigkeit über Standards anzuregen und die Leistungen

⁴¹ GEORG FEUSER, *Behinderte Kinder und Jugendliche: Zwischen Integration und Aussonderung*. Darmstadt 2005, 173 f.

der Schülerinnen und Schüler angemessen würdigen und fassen zu können. Dabei darf die Kritik einer Formalisierung von Bildung und Beurteilung nicht vergessen werden: Ein analytischer, kriterienorientierter Bewertungsansatz muss die Teilaspekte so ausgeglichen gewichten, dass der Gesamteindruck auf das Ergebnis nicht aus dem Blick gerät, indem entweder einzelne Aspekte unbegründet intensiv hervorgehoben werden (vgl. das Fallbeispiel in Kapitel 7.5.1, in welchem 30 der 60 in der Urteilsaufgabe erreichbaren Punkte durch Gestaltungsaspekte erreicht werden können), oder aber durch starre Kategorien der Offenheit der Aufgabe nicht gerecht werden. In Bezug auf letzteres ist beispielsweise an das Kriterium der Folgenreflexion zu denken (vgl. Kapitel 9.2.2), welches in Hinblick auf einige theologische Sachfragen in den Hintergrund rückt, für andere gesellschaftsrelevante ethische Fragen jedoch von äußerster Relevanz ist. Ebenso können Schülerlösungen Kategorien eröffnen, welche in einem analytischen Raster nicht vorgesehen sind. So spricht sich zum Beispiel JOHANNES ODENDAHL vor dem Hintergrund der Auseinandersetzung mit dem Zentralabitur im Fach Deutsch entschieden gegen ein Bewertungssystem nach Kriterienrastern aus.⁴² Seine Ansicht, es sei nicht möglich, die inhaltliche Qualität eines Deutsch-Aufsatzes lückenlos zu antizipieren und zu quantifizieren, lässt sich auf die in Bezug auf Offenheit und Komplexität sehr ähnlichen Urteilsbildungsaufgaben übertragen. Bewertungskataloge können seiner Meinung nach den eröffneten Freiräumen und der kreativen Entfaltung der Schülerinnen und Schüler nicht gerecht werden.⁴³ Er hebt in seinem Aufsatz hervor, dass das Urteil über eine Schülerleistung nicht vom Subjekt des Beurteilenden zu lösen ist und es ein »objektives« Urteil folglich nicht geben kann.⁴⁴

Interessante, spezifisch religionspädagogische Einwände in Bezug auf die Arbeit mit analytischen Kategorien arbeiten VIERA PIRKER und MARIA JUEN heraus, die in einer qualitativen Interviewstudie zwölf Religionslehrkräfte hinsichtlich ihrer Erfahrungen im Umgang mit einem Kompetenzraster zur Leistungsbewertung befragten.⁴⁵ Überwiegend wird die Arbeit mit dem Kompetenzraster, welches neben der sozialen und personalen Kompetenz sechs inhaltliche Kompetenzen benennt, jedoch keine Kompetenzniveaus differenziert, positiv wahrgenommen: Festgehalten wird von den Lehrkräften eine erhöhte Transparenz der Notengebung, eine aktivere Beteiligung der Schülerinnen und Schüler am Unterricht durch die offene Kommunikation über Bewertungskrite-

⁴² Vgl. JOHANNES ODENDAHL, Zentral, nicht zerebral. Wider die Benotung zentral gestellter Klausuren anhand eines Bewertungsrasters, in: Pädagogische Rundschau 62 (2008), 451–457.

⁴³ Vgl. a. a. O., 456.

⁴⁴ Vgl. ebd.

⁴⁵ Vgl. VIERA PIRKER/MARIA JUEN, Religion – (k)ein Fach wie jedes andere. Spannungsfelder und Perspektiven in der kompetenzorientierten Leistungsbeurteilung. Stuttgart 2018.

rien sowie eine verbesserte Wahrnehmung des Faches seitens Eltern, Kolleginnen und Kollegen.⁴⁶ Darüber hinaus wird eine Diskrepanz zwischen der eigenen Bewertungspraxis und der Bewertung mit den Kompetenzrastern konstatiert: Bei strikter Anwendung, so die Lehrerinnen und Lehrer, würden sich die Religionsnoten verschlechtern.⁴⁷ Die Studie PIRKER und JUENS zeigt, dass dies im Widerspruch zu ihrem Fachverständnis steht, welches »primär an der individuellen Bezugsnorm orientiert« ist und die »Ausnutzung der Notenspektrums von ›Sehr gut‹ bis ›Nicht genügend‹ verbietet« – demnach bestimmen Mildeeffekte und Großzügigkeit die Notengebung im Religionsunterricht.⁴⁸ Es wird die Angst formuliert, dass sich Schülerinnen und Schüler in Folge einer Umstellung auf eine kriterienorientierten Leistungsbewertung mittels des Kompetenzrasters vom Religionsunterricht abmelden könnten.⁴⁹ PIRKER und JÜEN leiten aus den Ergebnissen ihrer Studie ab, dass sich

»[r]eligionsunterrichtlich [...] eine stärkere Orientierung an Formen der Leistungsbeurteilung [empfiehlt], die vorrangig die pädagogische Funktion in den Blick nehmen, einzelne Schülerinnen und Schüler fördern und tendenziell auf überfachliche, nicht inhaltlich definierte Aspekte von Leistungen gerichtet sind.«⁵⁰

Sie plädieren für eine Akzentuierung der individuellen Bezugsnorm.⁵¹ Angesichts der von PIRKER und JÜEN erhobenen Diskrepanzen zwischen einer Bewertungspraxis mit und einer Bewertungspraxis ohne Analysekriterien und angesichts des von ihnen belegten Befundes, dass eine Arbeit mit Analysekriterien zu einer Reflexion und zur Professionalisierung der eigenen Bewertungspraxis führt, scheinen ihre Schlussfolgerungen gerade die formativen Aspekte der Leistungsbewertung zu betonen, welche ohnehin im Religionsun-

⁴⁶ Vgl. a.a.O., 192. Das von der Studie eruierte Kompetenzraster ist unter folgender Internetadresse einsehbar: https://blog.kohlhammer.de/wp-content/uploads/Zusatzmaterial_Religion_k-ein_Fach_wie_jedes_andere.pdf.

⁴⁷ Vgl. a.a.O., 193 und 190.

⁴⁸ Vgl. a.a.O., 196 f.

⁴⁹ Vgl. ebd.

⁵⁰ Vgl. a.a.O., 201.

⁵¹ Vgl. ebd. REIS und SCHWARZKOPF hingegen sprechen sich entschieden dagegen aus, die individuelle Bezugsnorm in den Diagnoseprozess zu integrieren, da sie ihr Diagnosemodell von Kompetenzerwartungen aus konzipieren (vgl. Kap. 2.3). Individuelle Maßstäbe halten sie ausschließlich auf der Ebene der Zeugnisnoten für sinnvoll. »Denn bei einer verantworteten Notenfindung für einen biografischen Abschnitt«, so REIS und SCHWARZKOPF, »spielt der individuelle Maßstab durchaus eine wichtige Rolle, aber nicht bei den punktuellen und isolierten Momenten der diagnostischen Leistungsmessung, die auf bestimmte, intentionale Lernziele ausgerichtet sind.« (vgl. DIES., Diagnose religiöser Lernprozesse, 45).

terricht schon einen breiten Anklang finden. Angesichts der in der vorliegenden Studie festgestellten mangelnden Sprachfähigkeit in Bezug auf messbare Bewertungskriterien ist im Gegenteil eine stärkere Diskussion über angemessene Analyseverfahren und eine modellgeleitete Diagnostik zu fordern. REIS und SCHWARZKOPF gehen in ihren Forderungen über PIRKER und JÜEN weit hinaus, wenn sie eine explizite Diagnosemethode für den Religionsunterricht fordern, die ausgehend von einer kompetenzorientierten Bildung von Lernzielen eine modellgeleitete Diagnostik ermöglicht (vgl. Kapitel 2.3).⁵² Sie warnen vor methodisch reduzierten Diagnoseformen, da diese »gefährlich in der Zuschreibung« seien und zu »didaktischen Fehlschlüssen« führten und fordern in diesem Zusammenhang, dass Diagnose ein eigenes professionelles Handlungsfeld wird und die notwendigen Ressourcen dafür erhält.⁵³ Dabei ist unbenommen, dass starre Standardisierungen weder komplexen Aufgabenstellungen noch den Lehrkräften und Prüflingen gerecht werden. So spricht sich DIETRICH BENNER gegen eine zu starke Standardisierung aus, welche den Raum für individuelle Lösungswege einschränkt. Er schreibt:

»Es widerspräche der Würde des Lehrers wie der Würde der Lernenden, wenn Prüfungen in Form von Kompetenztests abgenommen und Lernende damit auf Testpersonen reduziert würden, deren Leistungen nur noch nach standardisierten Vorgaben zu beurteilen sind.«⁵⁴

REIS und SCHWARZKOPF halten jedoch dagegen, dass eine Wertschätzung der Schülerinnen und Schüler gerade erst durch klar definierte Anforderungen Wirkung entfalten kann.⁵⁵ Sie schreiben:

⁵² Vgl. REIS/SCHWARZKOPF, Diagnosemodelle für den Religionsunterricht, 93 f.

⁵³ Vgl. a. a. O., 94 f.

⁵⁴ Vgl. DIETRICH BENNER, Pädagogische Auslegung der Anforderungsniveaus, in: DIETRICH BENNER/ROUMIANA NIKOLOVA (Hrsg.), Ethisch-moralische Kompetenz als Teil öffentlicher Bildung. Paderborn 2016, 147–158, hier 157.

Ähnlich kritisch gegenüber zu engen Lösungserwartungen äußern sich SCHÖNEMANN, THÜNEMANN und ZÜLSDORF-KERSTING in Bezug auf den Geschichtsunterricht, wenn sie für eine Deregulierung der Erwartungshorizonte plädieren: »Für die Formulierung der Erwartungshorizonte regen wir an, die Lösungserwartungen insbesondere im Anforderungsbereich III deutlich ergebnisoffener zu fassen. Dieses Plädoyer für eine konsequente De-Regulierung speist sich aus der Einsicht, dass gerade die anspruchsvollen historischen Lernleistungen sich kaum inhaltskonkret und vorab konkretisieren lassen« (vgl. DIES., Was können Abiturienten?, 123).

⁵⁵ Vgl. a. a. O., 95 sowie REIS/SCHWARZKOPF: Diagnose religiöser Lernprozesse, 44 f.

»Paradoerweise sorgt gerade das Unterlaufen einer komplexen Diagnostik dafür, dass die Lerngruppen künstlich homogenisiert werden (*alle* sind verschieden, kaum noch kirchlich sozialisiert, kein Interesse an theologischen Inhalten...). Subjektorientierung im religiösen Lernen braucht gerade explizite Normen [...], zu denen sich die Lernenden in ihrer Bildungsfreiheit verhalten können.«⁵⁶

Ohne eine Konkretisierung dessen, was erwartet wird – dies konnte die Analyse der Erwartungshorizonte zeigen – bleiben die Anforderungen unbestimmt und entschieden von den individuellen, domänenspezifischen Fähigkeiten der korrigierenden Lehrperson abhängig. Darüber hinaus entsteht Intransparenz, die keine Anknüpfungspunkte für den Diskurs über Qualitätsmerkmale bietet. Dennoch muss die formulierte Kritik insofern ernst genommen werden, als dass die Gesamtleistung nicht durch eine Zergliederung in Analyse Kriterien aus dem Blick gerät. Die von WERNER SACHER im Begriff der »übersummativen Qualität« gefassten Gesamtwirkung komplexer Leistungen wie Urteilsbildungen muss durch eine synthetische Leistungsbewertung, die von einem Gesamteindruck ausgeht, bewusst gesucht und wahrgenommen werden. Diese muss jedoch durch einen analytischen Blick Ergänzung und Prüfung finden, wie es WERNER SACHER in seinem hermeneutischen Kreisprozess der Bewertung vorschlägt (vgl. Kapitel 3.3.2).⁵⁷ Vor dem Hintergrund der Ergebnisse der Studie scheint ein Diskurs über die Qualitätsmerkmale von Urteilsfähigkeit zentral, die eine synthetische Bewertung durch analytische und in Niveaustufen differenzierende Kategorien zu ergänzen vermögen. Darüber hinaus kann durch die Kennzeichnung der Lösungsvorschläge als essenzielle oder optionale Bestandteile ein Mittelweg gefunden werden, der der korrigierenden Lehrkraft sowohl Orientierung als auch Freiraum in ihrer Bewertung geben kann.

9.3 Bekenntnisorientierung im Kontext Urteilsbildung

Als ein dritter diskussionswürdiger Aspekt erscheint vor dem Hintergrund der Ergebnisse der Studie die Rolle der Bekenntnisorientierung des Religionsunterrichts im Urteilsbildungsprozess. Anders als der Biologie-, Politik- oder Geographieunterricht hat der evangelische Religionsunterricht eine verfassungsrechtlich verankerte Bekenntnisorientierung und wird »in Übereinstimmung mit den Grundsätzen« der evangelischen Glaubensgemeinschaft erteilt (vgl. GG Art. 7.3 sowie Kapitel 4.2.1). Die Analyse hat ergeben, dass alle Aufgabenstel-

⁵⁶ REIS/SCHWARZKOPF, Diagnosemodelle für den Religionsunterricht, 95. REIS und SCHWARZKOPF verweisen bei dieser Ausführung auf JÜRGEN LINK, Versuch für den Normalismus. Wie Normalität produziert wird. Göttingen 1998, 52.

⁵⁷ SACHER, Leistungen entwickeln, überprüfen und beurteilen, 155.

lungen einen Bezug zur Religion herstellen, die Art dieser Bezugnahme jedoch unterschiedlich gestaltet sein kann: In einigen Fällen sind die Fragestellungen evangelischer, christlicher oder religiöser Art. In anderen Fällen sind es die Reflexionskriterien, welche eine christliche Perspektive auf das Thema einzunehmen fordern. Beide Arten von Aufgabenstellungen setzen einen gewissen Grad an Akzeptanz der Bedeutsamkeit theologischer Themen und Reflexionskategorien beim Prüfling voraus. Sie lassen sich jedoch hinsichtlich der Einnahme einer religiösen Innenperspektive voneinander unterscheiden, wie im Folgenden gezeigt werden soll.

9.3.1 Bekenntnisorientierung im Kontext theologischer Fragestellungen

Theologische Fragestellungen setzten vom Prüfling voraus, sich auf bekenntnisorientierte Prämissen einzulassen, ohne deren Annahme die Fragestellung nur verkürzt diskutiert werden könnte. So setzt beispielsweise eine Auseinandersetzung mit FULBERT STEFFENSKYS Gedanken über das Wesen Gottes und die Schuldhaftigkeit der Menschen, wie sie in einer Aufgabenstellung aus Niedersachsen im Jahr 2017 gefordert wurde [NI, 2017, gA (1) sowie NI, 2017, eA (2)] voraus, dass man sich auf die Prämisse einlässt, dass es einen Gott gibt, der sich mit den Menschen in eine Beziehung setzt. Eine Lösung der Aufgabe, die diese Prämisse in Frage stellt, würde in der Argumentation vorzeitig zu einem Ende führen, ohne den im Erwartungshorizont geforderten Grad an theologischer Reflexion zu erreichen, der durch den Anspruch zum Ausdruck kommt, dass der Prüfling die »Spannung von deus absconditus und deus revelatus« – also einem Ehrfurcht heischenden fremden Gott auf der einen und einem liebenden nahbaren Gott auf der anderen Seite – in seiner Auseinandersetzung berücksichtigt [E-NI, 2017, gA (1)].

Das Einlassen auf christliche Annahmen über die Welt macht die Bekenntnisorientierung des Faches aus und prägt die Urteilsbildung im Fach Evangelische Religionslehre. Mit der Bekenntnisorientierung verbunden ist die Freiwilligkeit des Religionsunterrichts: Das Einlassen auf die christliche Weltannahmen wird allen offen gestellt. Die Akzeptanz der Prämisse »Gott existiert« macht die Diskussion über theologische Fragestellungen in vielen Fällen erst möglich. Wesentlich ist, dass diese Prämisse deskriptiver und nicht normativer Art ist: Sie verlangt von Schülerinnen und Schülern keine Werthaltung und damit kein persönliches Bekenntnis. Sie können der Prämisse auf theoretischer Ebene – in

einem »Als-Ob-Modus« – folgen, ohne das Bekenntnis als »gut« oder »schlecht« bewerten zu müssen.⁵⁸

Ebenso bleibt die Offenheit der Urteilsbildung von der Akzeptanz der christlichen Prämissen unberührt: Die in einer kontroversen Fragestellung zur Diskussion stehenden Optionen sind nicht an ein Bekenntnis gebunden, sondern wie in jedem anderen Fach auch hinsichtlich des Urteilsentscheides offen. Erst das Vorhandensein konträrer, vertretbarer Optionen macht eine Urteilsbildung notwendig und sinnvoll (vgl. Kapitel 2.1). So geht es auch im Religionsunterricht nicht darum, eine spezifische Haltung, eine Überzeugung oder ein Bekenntnis zu prüfen, sondern den Zugriff auf die Kontroverse zu bewerten (vgl. zu den zu prüfenden Kriterien Kapitel 9.2); ganz abgesehen davon, dass es »die eine christliche Position« nicht gibt, die man würde prüfen können.

Die Analyse der Abituraufgabensets hat gezeigt, dass die benannte Offenheit an Deutungen und Urteilsentscheiden in einigen Erwartungshorizonten noch nicht eindeutig herausgestellt wird und dass einige Aussagen in ihrem Verbindlichkeitsgrad und in der Reflexion ihrer Tragweite noch eindeutiger als Deutungsoptionen markiert werden könnten (vgl. beispielsweise den in Kapitel 7.5.1 dargestellten Fall, in dem »Gottesliebe und Nächstenliebe als höchste ethisch-moralische Grundsätze für das Zusammenleben« gedeutet werden sollen, ohne das deutlich wird, ob der Prüfling diese normative Setzung zwangsläufig vertreten muss [E-TH, 2014 (2)]). Unbenommen davon ist, dass sowohl in den bildungsadministrativen Vorgaben als auch in vielen Erwartungshorizonten die Eigenständigkeit der Urteilsbildung betont wird, was auf die Offenheit der Urteilsbildung hinweist.

9.3.2 Bekenntnisorientierung in Bezug auf Reflexionskriterien

Eine zweite Gruppe an Aufgabenstellungen generiert den Bezug zum christlichen Glauben nicht über die Fragestellung, sondern über Reflexionskriterien. Im Untersuchungszeitraum wurden drei Kriterien mehr als einmal genannt: ein christliches Menschenbild, theologische Ethik und geeignete Bibeltexte (vgl. Kapitel 6.2.3). Anders als die deskriptiven Prämissen zu einer christlichen

⁵⁸ Die Spannung zwischen Teilnahme und Verstehen vs. Zustimmung und Bekenntnis hat die performative Religionsdidaktik intensiv bearbeitet und sowohl ein »Probehandeln« als auch ein »Verstehen des religiösen Weltzugangs« in Abgrenzung zur Übernahme von Bekenntnissen hervorgehoben. BERNHARD DRESSLER schreibt: »Teilnahme impliziert Positionalität. Aber die Berücksichtigung von Teilnahmeperspektiven darf nicht einfach als deren Übernahme verstanden werden. Verstehen bedeutet keinen Zwang zum Einverständnis« (vgl. BERNHARD DRESSLER, Art. Performativer Religionsunterricht, in: Wissenschaftliches Religionspädagogisches Lexikon im Internet [www.wirelex.de] 2015).

Weltsicht, die zunächst keine Wertung enthalten (beispielsweise »Gott existiert«), sind mit diesen Reflexionskriterien normative Setzungen verbunden: Wenn beispielsweise bei der Bewertung des Konzepts der Kryonik⁵⁹ Aspekte christlichen Menschenverständnisses und christlicher Zukunftshoffnung berücksichtigt werden sollen [BY, 2017 (4)], bilden diese zwei Kategorien, welche mit spezifischen Werten verbunden sind. So ist aus christlicher Anthropologie der Mensch ein endliches Geschöpf Gottes, welches seinen Sinn durch Gott zugesprochen bekommt und nicht durch sich selbst. Das Konzept der Kryonik muss mit diesen Grundprämissen in ein Verhältnis gebracht werden, wenngleich im Rahmen der christlichen Anthropologie verschiedene Akzentsetzungen möglich sind, die in Bezug auf die Aufgabe herausgearbeitet werden können. Die Grundprämissen der christlichen Anthropologie müssen dabei von den Schülerinnen und Schülern nicht übernommen, sie müssen aber gekannt und geprüft werden.

Auch in anderen Fachkulturen wurden verschiedene ethische Kriterien herausgearbeitet, die für domänenspezifische Urteilsbildungen leitend sind. PETER MASSING arbeitete für den Politikunterricht die Kategorie der *Effizienz für Sachurteile* heraus.⁶⁰ Diese fragt danach, mit welchen Mitteln bestimmte Zwecke erreicht werden können: also beispielsweise nach Handlungsmöglichkeiten, Aufwand und Kosten. Für Werturteile hebt er die Kategorie der *Legitimität* hervor.⁶¹ Sie muss sich unter anderem messen lassen an den Werten der Gesellschaft (Humanverträglichkeit), an dem Wohl der Betroffenen (Sozialverträglichkeit) und an den Folgen für die Umwelt (Umweltverträglichkeit).⁶²

KATRIN PLATZER und KRISTINA SINEMUS greifen in einem interdisziplinären Projekt zwischen Theologie und Biochemie am Beispiel einer bioethischen Fragestellung zum Anbau von gentechnisch verändertem Mais ähnliche Urteilkriterien auf.⁶³ Als Sachgerechtigkeitskriterien benennen sie Funktionsfähigkeit, Wirtschaftlichkeit und Sicherheit.⁶⁴ Unter ethischer Perspektive prüfen sie das Kriterium Humanorientierung, welches erfüllt ist, wenn der intendierte Einsatz oder unbeabsichtigte Folgen weder dem menschlichen Leben noch der menschlichen Identität und Integrität schaden bzw. wenn diese gefördert werden; die Umweltorientierung, die nach der Förderlichkeit für das Verhältnis von

⁵⁹ Wie bereits oben erwähnt, bezeichnet Kryonik ein Verfahren, bei dem Organismen oder einzelne Organe nach dem Tod durch Einfrieren in flüssigem Stickstoff konserviert werden, um sie zu einem späteren Zeitpunkt reanimieren zu können.

⁶⁰ Vgl. PETER MASSING, Kategorien des politischen Urteilens und Wege zur politischen Urteilsbildung, in: DERS./GEORG WEIBENO (Hrsg.), Politische Urteilsbildung, Schwalbach/Ts. 1997, 115–131, hier 124f.

⁶¹ Vgl. ebd.

⁶² Vgl. ebd.

⁶³ Vgl. PLATZER/SINEMUS, Ein Strukturmodell ethischer Urteilsbildung, 91–107.

⁶⁴ Vgl. a. a. O., 96–98.

Mensch, Umwelt und Technik fragt; die Sozialorientierung, die Schäden oder Nachteile für die menschliche Gesellschaft zu vermeiden sucht und die Zukunftsorientierung, welche die erstgenannten Kriterien hinsichtlich ihrer räumlichen und zeitlichen Fernwirkungen überprüft.⁶⁵

Auch wenn – wie oben bereits angemerkt – eine Prüfung der Kategorie nicht mit der Übernahme der zugrunde liegenden Werte einhergehen muss, enthalten die Kategorien doch normative Setzungen, welche auch in den bildungsadministrativen Vorgaben enthalten sind:

Die EPA im Fach Evangelische Religionslehre beispielsweise formulieren unter der Kompetenz Urteilsfähigkeit die Anforderung, »die Menschenwürde theologisch begründen und als Grundwert in aktuellen ethischen Konflikten zur Geltung bringen« zu können.⁶⁶ Ohne dass explizit gesagt wird, dass ein »Grundwert« höher zu schätzen ist als andere Werte und ob es weitere »Grundwerte« gibt, wird hier das normativ gesetzte Ziel formuliert, Schülerinnen und Schüler für Menschenwürde zu sensibilisieren.

Dass der Anspruch eines weltanschaulich-neutralen Staates nicht sein kann, überhaupt keine Werte zu vertreten, sondern dass es ihm darum geht, sich keiner spezifischen, Werte begründenden Weltdeutung anzuschließen, wurde bereits in Kapitel 4.2.1 deutlich. Vor diesem Hintergrund erklärt sich, dass auch bildungsadministrative Vorgaben generell sowie die Lehrpläne andere Fächer deutliche normative Setzungen vornehmen. Dass sich der Staat auch hinsichtlich schulischer Bildungs- und Erziehungsziele Werten verpflichtet, wird beispielsweise an folgendem Passus aus dem Thüringer Schulgesetz deutlich, welcher zahlreiche normative Setzungen enthält:

»[...] Die Schule erzieht zur Achtung vor dem menschlichen Leben, zur Verantwortung für die Gemeinschaft und zu einem verantwortlichen Umgang mit der Umwelt und der Natur. Sie pflegt die Verbundenheit mit der Heimat in Thüringen und in Deutschland, fördert die Offenheit gegenüber Europa und weckt das Verantwortungsgefühl für alle Menschen in der Welt. [...] Die Schülerinnen und Schüler lernen, ihre Beziehungen zu anderen Menschen nach den Grundsätzen der Gerechtigkeit, der Solidarität und der Toleranz sowie der Gleichberechtigung der Geschlechter zu gestalten. Dabei werden die Lernenden darauf vorbereitet, Aufgaben in Familie, Gesellschaft und Staat zu übernehmen und dazu angehalten, sich im Geiste des Humanismus und der christlichen Nächstenliebe für die Mitmenschen einzusetzen.«⁶⁷

⁶⁵ Vgl. a. a. O., 98–103.

⁶⁶ EPA, 9.

⁶⁷ Vgl. THÜRINGER SCHULGESETZ vom 6. August 1993 (GVBl. 445) in der Fassung der Bekanntmachung vom 30. April 2003 (GVBl. 238) zuletzt geändert durch Artikel 5 des Gesetzes vom 31. Januar 2013 (GVBl. 22, 23).

Ein weiteres Beispiel wertorientierter Setzungen findet sich im baden-württembergischen Lehrplan für Geographie, in welchem das Ziel formuliert ist, dass die Schülerinnen und Schüler »die Bedeutung des Nachhaltigkeitsprinzips für den Erhalt der Kultur, für die Leistungsfähigkeit der Wirtschaft, für die gesellschaftlichen Bedürfnisse sowie die Entwicklungs- und Lebenschancen zukünftiger Generationen«⁶⁸ erkennen. Die Sensibilisierung für Nachhaltigkeit und der Aufbau einer entsprechenden Haltung seitens der Schülerinnen und Schüler ist hier als die Aufgabe des Faches deutlich benannt.⁶⁹ Und auch in den Bildungsstandards des Faches Biologie für den Mittleren Schulabschluss, wird in Bezug auf Urteilskompetenz formuliert, dass »Schülerinnen und Schüler [...] Wertschätzung für eine intakte Natur und eine eigene gesunde Lebensführung« entwickeln sowie »Verständnis für Entscheidungen im Sinne einer nachhaltigen Entwicklung« zeigen.⁷⁰

Urteilsbildungen benötigen diese Kriterien. In den Aufgabenstellungen, in denen die Kriterien der Urteilsbildung nicht vorgegeben sind, müssen Schülerinnen und Schüler selbst Kriterien finden und diese besten Falls explizit ausweisen. So können die Aufgabenstellungen

»3. Beurteilen Sie, ob Missalla überzeugende Antworten auf religionskritische Einwände gibt.« [NI 2018 gA (2)] oder

»4. Setzen Sie sich mit Mettes Aussagen über die Versuchung (Z.28 bis Ende) auseinander.« [NI 2019 eA (2)]

nur beantwortet werden, wenn man einen Maßstab hat, an welchem die Aussagen geprüft werden. Die Erwartungshorizonte der Aufgabenstellungen führen diese Kriterien jedoch nicht aus, sondern bringen in beiden Fällen inhaltliche Pro- und Kontraargumente an. Das Abstraktionsniveau, welches in der Offenlegung der

⁶⁸ MINISTERIUM FÜR KULTUS, BILDUNG UND SPORT BADEN-WÜRTTEMBERG (Hrsg.), Bildungsplan des Gymnasiums. Geographie. Stuttgart 2016, 5. Abrufbar unter: http://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lsw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_GYM_GEO.pdf (03.06.2020).

⁶⁹ Interessant ist in diesem Zusammenhang eine von FELZMANN zitierte Aufgabenstellung aus dem schriftlichen Abitur im Fach Geographie aus Baden-Württemberg, in welcher Nachhaltigkeit als ein Reflexionskriterium genannt wird: »In Bremerhaven wurde mit dem Container-Terminal 4 im September 2008 eines der bislang größten maritimen Bauvorhaben Europas fertiggestellt. Bewerten Sie diese Hafenerweiterung im Bereich des ostfriesischen Wattmeeres unter dem Aspekt der Nachhaltigkeit.« (vgl. Aufgabenstellung aus der schriftlichen Abiturprüfung im Fach Geographie aus Baden-Württemberg 2009. Zitiert nach FELZMANN, Das Bewerten bewerten, 19.)

⁷⁰ Vgl. STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER (Hrsg.), Bildungsstandards Biologie, 12.

Kriterien und Wertmaßstäbe im Forschungskontext gefordert wird, wird hier nicht von den Erwartungshorizonten eingeholt (vgl. Kapitel 4 sowie Kapitel 9.2.2 zur Transparenz der zugrunde liegenden Werte). Dennoch werden zwei konträre Positionen genannt. Der Prüfling hat die Möglichkeit sich zu entscheiden. Die Offenheit der Positionierung ist folglich gegeben. Was jedoch der Zustimmung bedarf ist die Relevanz der Fragestellung und der Urteilskriterien, wobei sich diese je nach Stärke des religiösen Bezugs unterscheiden: Je näher eine Fragestellung die Kernbereiche christlicher Theologie berührt, umso stärker erfordert sie eine Zustimmung in Bezug auf die Relevanz: Teilt man die christliche Weltansicht nicht, so ist Mettes Interpretation der Versuchung irrelevant.

DIRK FELZMANN konstatiert in Bezug auf Reflexionskriterien im schriftlichen Abitur des Faches Geographie:

»Da der Umfang, in dem Kriterien für das Urteilfällen vorgegeben sind, differiert, liegt die kognitive Herausforderung auf Seiten der Schüler einmal mehr beim plausiblen und transparenten Erstellen eigenständiger Kriterienkataloge, einmal mehr beim inhaltlich angemessenen Transfer erlernter Kriterien (z.B. »nachhaltige Entwicklung«) auf die konkrete Situation.«⁷¹

Dabei ist zentral, dass nicht alle Kriterien in allen ethischen Fragestellungen gleichermaßen Relevanz besitzen, so dass die Kriterien auf jede Frage neu angepasst werden müssen. In Bezug auf metaphysisch-religiöse Fragen sind Bewertungskriterien wie die Sozial-, Umwelt- und Humanorientierung oftmals sekundär. Sie sind in andere Argumentationsmuster gestellt, welche sich von einer spezifischen Weltansicht und einer göttlichen Offenbarung her begründen. Die in den Erwartungshorizonten genannten Reflexionskriterien eines christlichen Menschenbildes, einer theologischen Ethik und biblischer Texte sind als ein Spezifikum religiöser Urteilsbildung zu fassen. Sie lassen sich nicht nur auf religiöse Fragestellungen, sondern auch auf allgemeinethische Fragestellungen anwenden, wenngleich bei diesen die genannten ethischen Kategorien ebenso mitbedacht sein müssen.

Ethische und religiöse Urteilsbildung im Religionsunterricht hat damit zwei Spezifika: Zum einen die Akzeptanz christlicher Prämissen in Bezug auf Wirklichkeitsannahmen und zum anderen theologische Reflexionskriterien, an denen Urteilsentscheide geprüft werden. Vor dem Hintergrund der analysierten Erwartungshorizonte kann der Wunsch formuliert werden, dass die Offenheit der Urteilsbildung – gerade vor der Bekenntnisorientierung des Faches – an vielen Stellen noch deutlicher herausgestellt wird.

⁷¹ FELZMANN, Das Bewerten bewerten, 19.

10 Résumé und Ausblick

Die Studie war motiviert von der Fragestellung, welche Qualitätsmerkmale sich für die Kompetenz Urteilsfähigkeit sowohl im Forschungsdiskurs als auch in schriftlichen Abituraufgabenstellungen und deren Erwartungshorizonten finden lassen, die eine Bewertung ermöglichen. Es konnte gezeigt werden, dass im Forschungsdiskurs unterschiedliche Qualitätsmerkmale diskutiert und Niveau-stufenmodelle erstellt und empirisch geprüft wurden. Die in der Forschung diskutierten Merkmale haben jedoch bisher nur in geringem Umfang Eingang in die Erwartungshorizonte der Abiturprüfungsaufgaben gefunden. Möchte man die Kompetenzorientierung ernst nehmen, so zeigen die Ergebnisse der Studie, dass die aktuellen Erwartungshorizonte den Ansprüchen eines kompetenzorientierten Prüfungsverständnisses noch nicht genügen¹: Der Anspruch eines kompetenzorientierten Religionsunterrichts ist nicht allein durch didaktische Leitprinzipien wie beispielsweise Metakognition, lebensweltlicher Anwendungsbezug oder Wissensvernetzung einzulösen – ähnliche Forderungen haben schon vorangegangene didaktische Konzeptionen erhoben. Die Ergebnisse der Studie haben gezeigt, dass auch die inhaltliche Ausformulierung der Kompetenzerwartungen weiter vorangetrieben werden muss. Religionsunterricht kann sich erst dann kompetenzorientiert nennen, wenn er eine klare Vorstellung davon hat, was die Qualität der anzustrebenden Kompetenzen ausmacht und wie ein stufenweiser Lernzuwachs in Bezug auf jede einzelne Kompetenz aussieht,

¹ »Schwammige Formulierungen« in Bezug auf Kompetenzbenennungen problematisiert auch MIRJAM ZIMMERMANN in einer Reflexion über den Ertrag der Kompetenzorientierung für den Religionsunterricht. Sie schreibt: »[...] Außerdem fragen sich die Studierenden schon beim oberflächlichen Blick in die Kerncurricula ihrer Schulform, wie man denn solche schwammigen Formulierungen z.B. zur biblischen Textauslegungskompetenz (zentral) kontrollieren könne. Sie sind deshalb bei der eigenständigen Planung oft unsicher, wie sie sinnvoll vorgehen können.« (vgl. HARTMUT LENHARD/MANFRED PIRNER/MIRJAM ZIMMERMANN, (Was) Hat der RU durch die Kompetenzorientierung gewonnen? Ein Schreibgespräch, in: Zeitschrift für Pädagogik und Theologie 70 (2018), 339–346, hier 341).

gezielt gefördert und valide geprüft werden kann. Die vagen Formulierungen in den Erwartungshorizonten über Qualitätsmerkmale korrespondieren mit den spezifischen Anforderungen der Urteilsbildung derzeit noch nicht. Es bedarf klarerer Vorstellungen darüber, was im Bereich Urteilsbildung erlernt und geprüft werden soll. Dabei wäre auch explizit zu benennen, was nicht bewertet werden sollte. Eine Verankerung von Niveaustufenmodellen zur Urteilsbildung in den Einheitlichen Prüfungsanforderungen, die verschiedene Ausprägungsgrade der Qualitätsmerkmale zu erfassen vermögen, könnte sich zum einen positiv auf die Berücksichtigung von Urteilskompetenz in den Erwartungshorizonten auswirken, da bisher in vielen der untersuchten Lösungsskizzen die geforderte Urteilsbildung nur unzureichende Beachtung findet. Zum anderen könnte sie sich positiv auf die Vergleichbarkeit der Anforderungen und letztendlich auch auf die Qualität des vorangehenden Unterrichts auswirken, indem einheitliche Standards geschaffen werden.

Die vorliegende Studie legt den Fokus auf bildungsadministrative Vorgaben. Die in den Erwartungshorizonten benannten Qualitätsmerkmale können als Diskursbeiträge zum Umgang mit kompetenzorientierten Anforderungen gelesen werden. Mit einer solchen Lesart ergibt sich der Eindruck, dass viele der gemachten Aussagen so formuliert sind, dass sie maximal offen »nach oben hin« sind und alle Leistungsfacetten zulassen, wobei auch ein niedriges Leistungsniveau nicht völlig aus dem Rahmen fällt. Es wäre daher ein lohnenswerter Gegenstand weiterer Forschung, mit welchen Kriterien Lehrkräfte bei der Korrektur von Urteilsbildung offene Formulierungen wie die einer »kritischen Stellungnahme« oder einer »schlüssigen und differenzierten Begründung« [E-BY, 2015 (4)] füllen. Inspirierend könnte dafür eine qualitative Interviewstudie aus der Biologiedidaktik von KERSTIN HEUSINGER VON WALDEGGE sein, die nach den Sichtweisen von Lehrkräften zur Beurteilung von Urteilskompetenz (in den Bildungsstandards Biologie als »Bewertungskompetenz« bezeichnet), nach ihren Kriterien der Beurteilung und nach Fördermaßnahmen (Methoden und Handlungsstrategien) der Lehrkräfte fragt.²

² Vgl. KERSTIN HEUSINGER VON WALDEGGE, Biologielehrkräfte diagnostizieren die Schülerkompetenz »Bewerten«. Eine qualitative Untersuchung zur Orientierung bei der Diagnose. Hamburg 2016.

Es handelt sich bei den acht für die Studie Interviewten um Lehrkräfte, die zuvor eine einjährige Weiterbildung zu Urteilskompetenz durchlaufen haben. Die Fokussierung auf »Expertinnen« und »Experten« dieser Kompetenz wird damit begründet, dass die Interviews eine Akzeptanz der Kompetenz im Fach Biologie voraussetzen, die bei Biologielehrkräften angesichts der eher jungen Verankerung der Kompetenz nicht unbedingt gegeben ist (vgl. Kap. 4.1.4). Eine Studie über die Einstellungen von Biologielehrkräften zur Urteilskompetenz von NELE ALFS hat ergeben, dass Urteilsfähigkeit bei naturwissenschaftlich ausgerichteten Lehrkräften als eine Kompetenz betrachtet wird, die eher gesellschaftswissenschaftlichen

Eine weitere sich anschließende Forschungsfrage betrifft den Umgang der Schülerinnen und Schüler mit Urteilsbildungsaufgaben. So ist bisher wenig erforscht, welche schwierigkeitskonstituierenden Merkmale zu Verstehenshürden bei Schülerinnen und Schülern führen. Durch die Analyse von Schülerinnen- und Schülerantworten könnte eine weitere fruchtbare Perspektive auf die Fragestellung nach Bewertungskriterien für Urteilsbildungsaufgaben gewonnen werden.

Fächern zuzuordnen ist und daher »fächerübergreifend mit den Fächern Religion, Werte und Normen, Politik und Deutsch stattfinden« sollte (vgl. ALFS et al., Bewertungsprozesse verstehen, 104). In der Außenwahrnehmung wird Urteilsfähigkeit folglich (auch) dem Religionsunterricht als Expertise zugeschrieben. Auch diese Beobachtung lässt eine Anschlussstudie über die Bewertungskriterien von Urteilsfähigkeit von Religionslehrkräften interessant erscheinen.

Literaturverzeichnis

- ALFS, NELE/HEUSINGER VON WALDEGGE, KERSTIN/HÖBLE, CORINNA: Bewertungsprozesse verstehen und diagnostizieren, in: Zeitschrift für interpretative Schul- und Unterrichtsforschung 1 (2012), 83–112.
- ALTRICHTER, HERBERT/MAAG MERKI, KATHARINA: Steuerung der Entwicklung des Schulwesens, in: DIES. (Hrsg.), Handbuch. Neue Steuerung im Schulwesen. Wiesbaden ²2016, 1–27.
- ANDERSEN, SVEND: Einführung in die Ethik. Berlin/New York ²2005.
- BAUMGARTNER, CHRISTOPH: Umweltethik – Umwelthandeln. Ein Beitrag zur Lösung des Motivationsproblems. Paderborn 2005.
- BAYERISCHES STAATSMINISTERIUM FÜR UNTERRICHT UND KULTUS (Hrsg.), Schulordnung für die Gymnasien in Bayern (Gymnasialschulordnung – GSO) vom 23. Januar 2007.
- BELJAN, JENS: Schule als Resonanzraum und Entfremdungszone. Eine neue Perspektive auf Bildung. Weinheim/Basel ²2019.
- BENNER, DIETRICH: Pädagogische Auslegung der Anforderungsniveaus, in: DIETRICH BENNER/ROUMIANA NIKOLOVA (Hrsg.), Ethisch-moralische Kompetenz als Teil öffentlicher Bildung. Paderborn 2016, 147–158.
- BENNER, DIETRICH: Allgemeine Pädagogik. Eine systematisch-problemgeschichtliche Einführung in die Grundstruktur pädagogischen Denkens und Handelns. Weinheim/Basel ⁸2015.
- BENNER, DIETRICH/DEGHANI, SHAMSI/NIKOLOVA, ROUMIANA/SCHARREL, JOANNA/SCHIEDER, ROLF/SCHLUß, HENNING/WEIß, THOMAS/WILLEMS, JOACHIM: Modellierung und Testung religiöser und ethischer Kompetenzen im Interesse ihrer Vergleichbarkeit, in: Zeitschrift für Pädagogik und Theologie 2 (2010), 165–174.
- BENNER, DIETRICH/IVANOV, STANISLAV/NIKOLOVA, ROUMIANA: Die Anforderungsniveaus der ethisch-moralischen Teilkompetenzen, in: DIETRICH BENNER/ROUMIANA NIKOLOVA (Hrsg.), Ethisch-moralische Kompetenz als Teil öffentlicher Bildung. Paderborn 2016, 131–146.
- BENNER, DIETRICH/ NIKOLOVA, ROUMIANA: Der Berliner Ansatz zur Konstruktion und Erhebung ethisch-moralischer Kompetenzniveaus im öffentlichen Erziehungs- und Bildungssystem, in: DIES. (Hrsg.), Ethisch-moralische Kompetenz als Teil öffentlicher Bildung. Paderborn 2016, 13–68.

- BERTELSMANN STIFTUNG (Hrsg.), Chancenspiegel – eine Zwischenbilanz. Zur Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme seit 2002. Gütersloh 2017.
- BIEHL, PETER/ROTHGANGEL, MARTIN: Konzeptionen und Strukturen, in: MICHAEL WERMKE/GOTTFRIED ADAM/MARTIN ROTHGANGEL (Hrsg.), Religion in der Sekundarstufe II. Ein Kompendium. Göttingen 2006, 183–218.
- BIELEFELDT, HEINER: Muslimische Minderheiten im säkularen Rechtsstaat, in: GERHARD KRIUP/WOLFGANG VÖGELE (Hrsg.), Schatten der Differenz. Das Paradigma der Anerkennung und die Realität gesellschaftlicher Konflikte. Hamburg 2006, 69–90.
- BOHL, THORSTEN: Schulische Notengebung: Probleme und Entwicklungsmöglichkeiten, in: Katechetische Blätter 132 (2007), 249–254.
- BOHNER, GERD: Einstellungen, in: WOLFGANG STROEBE/KLAUS JONAS/MILES HEWSTONE (Hrsg.), Sozialpsychologie. Eine Einführung. Berlin u. a. ⁴2002, 266–314.
- BÜCHTER, ANDREAS/ PALLACK, ANDREAS: Methodische Überlegungen und empirische Analysen zur impliziten Standardsetzung durch zentrale Prüfungen, in: Journal für Mathematik-Didaktik 33 (2012), 59–85.
- BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (Hrsg.), Zur Entwicklung nationaler Bildungsstandards. Expertise. Bonn/Berlin 2007. Abrufbar unter: https://www.researchgate.net/publication/281345382_Zur_Entwicklung_nationaler_Bildungsstandards_Eine_Expertise_Stand_Juni_2003 (28.8.2020).
- BÜTTNER, GERHARD/DIETRICH, VEIT-JAKOBUS: Entwicklungspsychologie in der Religionspädagogik. Göttingen 2016.
- CHENG, LIYING/CURTIS, ANDY: Test impact and washback: implications for teaching and learning, in: CHRISTINE COOMBE/BARRY O’SULLIVAN/PETER DAVIDSON/STEVEN STOYN-OFF (eds.): Cambridge Guide to Second Language Assessment. Cambridge 2012, 89–95.
- CHRISTMANN, URSULA/GROEBEN, NORBERT/SCHREIER, MARGRIT: Subjektive Theorien – Rekonstruktion und Dialog-Konsens, in: SPIEL 18 (1999), 138–154.
- COLARDYN, DANIELLE/BJORNAVOLD, JENS: Validation of Formal, Non-Formal and Informal Learning. Policy and practices in EU Member States, in: European Journal of Education 39 (2004), 69–89.
- CRAMER, CAROLIN: Chancen und Grenzen der Kompetenzorientierung in Religionsunterricht und Religionslehrerbildung, in: Zeitschrift für Pädagogik und Theologie 70 (2018), 372–385.
- DETJEN, JOACHIM/MASSING, PETER/RICHTER, DAGMAR/WEIBENO, GEORG: Politikkompetenz – ein Modell. Wiesbaden 2012.
- DICKEL, MIRKA: Ethisches Können. Geographieunterricht als Ethisches Können, in: MIRKA DICKEL/ANKE JOHN/MICHAEL MAY/KATHARINA MUTH/LAURENZ VOLKMAN/MARIO ZIEGLER (Hrsg.), Urteilspraxis und Wertmaßstäbe im Unterricht. Ethik, Englisch, Geschichte, politische Bildung und Religion. Schwalbach 2020, 71–99.
- DIETERICH, VEIT-JAKOBUS: Wie steht es mit der Kompetenzorientierung in den Lehr- und Bildungsplänen für den Religionsunterricht? Eine Standortbestimmung nach ein- einhalb Jahrzehnten Curriculumrevision, in: Zeitschrift für Pädagogik und Theologie 70 (2018), 347–359.
- DISDORN-LIESEN, VARVARA: Vergleichbarkeit in der Vielfalt. Leistungsanforderungen und Leistungsfeststellungen im Zentralabitur Deutsch. Wiesbaden 2016.

- DRESSLER, BERNHARD: Art. Performativer Religionsunterricht, in: Wissenschaftliches Religionspädagogisches Lexikon im Internet (www.wirelex.de) 2015.
- DRESSLER, BERNHARD: Religionsunterricht – mehr Kompetenzorientierung? In: ANDREAS FEINDT/VÖLKER EISENBAST/PETER SCHREINER/ALBRECHT SCHÖLL (Hrsg.), Kompetenzorientierung im Religionsunterricht. Befunde und Perspektiven. Münster u. a. 2009, 23–37.
- DRESSLER, BERNHARD: Unterscheidungen. Religion und Bildung. Leipzig 2006.
- DRESSLER, BERNHARD: Religionsunterricht als Werteerziehung? Eine Problemanzeige, in: Zeitschrift für Evangelische Ethik 46 (2002), 256–269.
- DUDENREDAKTION (Hrsg.), Art. Philosophie, in: Duden online. www.duden.de/rechtschreibung/Philosophie (15.09.2020).
- EID, MICHAEL/GOLLWITZER, MARIO/SCHMITT, MANFRED: Statistik und Forschungsmethoden. Weinheim/Basel³ 2017.
- EISENBAST, VÖLKER/FISCHER, DIETLIND (Hrsg.), Stellungnahmen und Kommentare zu »Grundlegende Kompetenzen religiöser Bildung«. Münster 2007.
- ENGLERT, RUDOLF: Die verschiedenen Komponenten ethischen Lernens und ihr Zusammenspiel. Überlegungen zu einem Gesamtprogramm ethischer Bildung, in: RUDOLF ENGLERT/HELGA KOHLER-SPIEGEL/ELISABETH NAURATH/BERND SCHRÖDER/FRIEDRICH SCHWEITZER (Hrsg.), Ethisches Lernen. Neukirchen-Vluyn 2015 [= Jahrbuch der Religionspädagogik 31], 108–118.
- ENGLERT, RUDOLF: Religionspädagogik in der Schule, in: FRIEDRICH SCHWEITZER/THOMAS SCHLAG (Hrsg.), Religionspädagogik im 21. Jahrhundert. Gütersloh 2004, 79–93.
- ERLER, MICHAEL: Philosophie I, in: Religion in Geschichte und Gegenwart⁴ 2003 online. http://dx-1doi-1org-1002dcb8v0557.han.ulb.uni-jena.de/10.1163/2405-8262_rgg4_COM_024378 (15.9.2020).
- EVANGELISCHE KIRCHE IN DEUTSCHLAND (Hrsg.), Religiöse Orientierung gewinnen. Evangelischer Religionsunterricht als Beitrag zu einer pluralitätsfähigen Schule. Gütersloh 2014.
- EVANGELISCHE KIRCHE IN DEUTSCHLAND (Hrsg.), Kompetenzen und Standards für den Evangelischen Religionsunterricht der Sekundarstufe I. Ein Orientierungsrahmen. Texte 111. Hannover 2011.
- EVANGELISCHE KIRCHE IN DEUTSCHLAND (Hrsg.), Kerncurriculum für das Fach Evangelische Religionslehre in der gymnasialen Oberstufe. Themen und Inhalte für die Entwicklung von Kompetenzen religiöser Bildung. Texte 109. Hannover 2010.
- EVANGELISCHE KIRCHE IN DEUTSCHLAND (Hrsg.), Theologisch-Religionspädagogische Kompetenz. Professionelle Kompetenzen und Standards für die Religionslehrausbildung. Empfehlungen der Gemischten Kommission zur Reform des Theologiestudiums. Hannover 2008.
- EVANGELISCHE KIRCHE IN DEUTSCHLAND (Hrsg.), Identität und Verständigung. Standort und Perspektiven des Religionsunterrichts in der Pluralität. Eine Denkschrift der EKD. Gütersloh 1994.
- FACHVERBAND PHILOSOPHIE E.V./FACHVERBANDS ETHIK E.V./FORUM FÜR DIDAKTIK DER PHILOSOPHIE UND ETHIK (Hrsg.), Dresdner Konsens für den Philosophie- und Ethikunterricht, in: Zeitschrift für Didaktik der Philosophie und Ethik 38 (2016), 106. Ab-

- rufbar unter: https://philosophiedidaktik.files.wordpress.com/2017/03/dresdner_konsens.pdf (16.05.2018).
- FELZMANN, DIRK: Das Bewerten bewerten. Klausuraufgaben zur Beurteilungs- und Bewertungskompetenz, in: *Geografie und Schule*, 35 (2013) 206, 15–22.
- FEND, HELMUT: Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen. Wiesbaden ²2008.
- FEUSER, GEORG: Behinderte Kinder und Jugendliche: Zwischen Integration und Aussonderung. Darmstadt ²2005.
- FISCHER, DIETLIND/ELSENBAST VÖLKER (Redaktion): Grundlegende Kompetenzen religiöser Bildung. Zur Entwicklung des evangelischen Religionsunterrichts durch Bildungsstandards für den Abschluss der Sekundarstufe I. Münster 2006.
- FOERSTER, HEINZ VON: Lethologie. Eine Theorie des Erlernens und Erwissens angesichts von Unwißbarem, Unbestimmbarem und Unentscheidbarem, in: REINHARD VOB (Hrsg.), *Die Schule neu erfinden. Systematisch-konstruktivistische Annäherungen an Schule und Pädagogik*. Neuwied ³1999.
- FUCHS, MONIKA E.: »Denn irgendwie krank sein bedeutet nicht schlechter sein.« – Bioethik aus Schülerperspektive am Beispiel Pränataldiagnostik, in: *Zeitschrift für interpretative Schul- und Unterrichtsforschung* 1 (2012), 184–199.
- FUCHS, MONIKA E.: Bioethische Urteilsbildung im Religionsunterricht. Theoretische Reflexion – Empirische Rekonstruktion. Göttingen 2010.
- FUCHS, MONIKA E.: Empirische und didaktische Perspektiven bioethischen Lernens im Religionsunterricht, in: *Theo-Web. Zeitschrift für Religionspädagogik* 8 (2009), 80–90.
- GLOY, HORST: Themen statt Texte? In: NORBERT SCHNEIDER (Hrsg.), *Religionsunterricht. Konflikte und Konzepte. Beiträge zu einer neuen Praxis*. Hamburg 1971, 67–79.
- GRÄSEL, CORNELIA/PARCHMANN, ILKA: Implementationsforschung – oder: der steinige Weg, Unterricht zu verändern, in: *Unterrichtswissenschaft* 32 (2004), 196–214.
- HAHN, MATTHIAS: Problemorientierter Religionsunterricht, in: *Wissenschaftliches Religionspädagogisches Lexikon im Internet* (www.wirelex.de) 2016.
- HAIDT, JONATHAN: The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment, in: *Psychological Review* 108 (2001), 814–834. Abrufbar unter: https://www.motherjones.com/wp-content/uploads/emotional_dog_and_rational_tail.pdf (09.08.2018).
- HANNKEN-ILLJES, KATI: *Argumentation. Einführung in die Theorie und Analyse der Argumentation*. Tübingen 2018.
- HELBING, DOMINIK: Religiöse Herausforderung und religiöse Kompetenz. Empirische Sondierungen zu einer subjektorientierten und kompetenzbasierten Religionsdidaktik. Wien/Zürich 2010.
- HEMEL, ULRICH: *Ziele religiöser Erziehung. Beiträge zu einer integrativen Theorie*. Frankfurt a.M. 1988.
- HENCKMANN, WOLFHART/KREB, HARTMUT: Art. Wert, in: *Theologische Realenzyklopädie* Bd. 35, 2003, 648–657.
- HERMS, EILERT: Art. Moral, in: *Religion in Geschichte und Gegenwart* ⁴2002 online. http://dx.doi.org/002dcbqo0d07.han.ulb.uni-jena.de/10.1163/2405-8262_rgg4_SIM_14450 (15.09.2020).

- HERMS, EILERT: Art. Ethik V, in: Religion in Geschichte und Gegenwart ⁴1999 online. http://dx-1doi-1org-1002dcb8v0557.han.ulb.uni-jena.de/10.1163/2405-8262_rgg4_COM_04670 (19.09.2020).
- HEYNITZ, MARTINA VON: Die Entwicklung von Testaufgaben zur Erhebung ethisch-moralischer Grundkenntnisse, Urteils- und Handlungsentwurfskompetenz, in: DIETRICH BENNER/ROUMIANA NIKOLOVA (Hrsg.), Ethisch-Moralische Kompetenz als Teil öffentlicher Bildung. Der Berliner Ansatz zur Konstruktion und Erhebung ethisch-moralischer Kompetenzniveaus im öffentlichen Erziehungs- und Bildungssystem mit einem Ausblick auf Projekte zu ETiK-International. Paderborn 2016, 69–100.
- HILDEBRANDT, UTA: Das Grundrecht auf Religionsunterricht, Tübingen 2000.
- HILGER, GEORG: Welche Wirkung hat der Religionsunterricht? Evaluation und Leistungsbewertung, in: GEORG HILGER/STEPHAN LEIMGRUBER/HANS-GEORG ZIEBERTZ (Hrsg.), Religionsdidaktik. Ein Leitfaden für das Studium, Ausbildung und Beruf. München ⁶2003, 282–290.
- HÖBLE, CORINNA: Moralische Urteilsfähigkeit. Eine Interventionsstudie zur moralischen Urteilsfähigkeit von Schülern zum Thema Gentechnik. Innsbruck u. a. 2001.
- HÖBLE, CORINNA/ HEUSINGER VON WALDEGGE, KERSTIN: Bewertungskompetenz diagnostizieren, in: Der mathematische und naturwissenschaftliche Unterricht 63 (2010), 428–434.
- HÜBNER, DIETMAR: Einführung in die Philosophische Ethik. Göttingen 2014.
- HUMBOLDT, WILHELM VON: Theorie der Bildung des Menschen, in: ANDREAS FLINTNER/KLAUS GIEL (Hrsg.), Wilhelm von Humboldt. Werke in fünf Bänden. Erster Band: Schriften zur Anthropologie und Geschichte. Berlin 1960.
- IVANOV, STANISLAV: Die Validierung des ETiK-Instruments, in: DIETRICH BENNER/ROUMIANA NIKOLOVA (Hrsg.), Ethisch-moralische Kompetenz als Teil öffentlicher Bildung. Paderborn 2016, 101–129.
- JECHLE, THOMAS: Kommunikatives Schreiben. Prozeß und Entwicklung aus der Sicht kognitiver Schreibforschung. Tübingen 1992.
- JEISMANN, KARL-ERNST: Didaktik der Geschichte: Das spezifische Bedingungsfeld des Geschichtsunterrichts, in: GÜNTER C. BEHRMANN/KARL-ERNST JEISMANN/HANS SÜSMUTH (Hrsg.), Geschichte und Politik. Didaktische Überlegungen eines kooperativen Unterrichts. Paderborn 1978, 50–107.
- JENDORFF, BERNHARD: Leistungsmessung im Religionsunterricht. Methoden und Beispiele. München 1979.
- JORDAN, ALEXANDER/ROSS, NATHALIE/KRAUSS, STEFAN/BAUMERT, JÜRGEN/BLUM, WERNER/NEUBRANDT, MICHAEL/LÖWEN, KATRIN/ BRUNNER, MARTIN/KUNTER, MAREIKE: Klassifikationsschema für Mathematikaufgaben. Dokumentation der Aufgabenkategorisierung im COACTIV-Projekt. Berlin 2006.
- JOST, JÖRG/BÖTTCHER, INGRID: Leistungen messen, bewerten und beurteilen, in: MICHAEL BECKER-MROTZEK/INGRID BÖTTCHER (Hrsg.), Schreibkompetenz entwickeln und beurteilen. Berlin ⁵2014, 113–144.
- KABISCH, RICHARD: Wie lehren wir Religion? Versuch einer Methodik des evangelischen Religionsunterrichts für alle Schulen auf psychologischer Grundlage. Göttingen ⁵1920.

- KAUFHOLD, MARISA: Kompetenz und Kompetenzerfassung. Analyse und Beurteilung von Verfahren der Kompetenzerfassung. Wiesbaden 2006.
- KAUFMANN, HANS BERNHARD: Muß die Bibel im Mittelpunkt des Religionsunterrichts stehen? In: FOLKERT RICKERS/BERNHARD DRESSLER (Hrsg.), Thematisch-problemorientierter Religionsunterricht. Aufbruch – Bewährung in der Praxis – Impulse, Neukirchen-Vluyn 2003, 145–151.
- KAYSER, JÖRG/HAGEMANN, ULRICH: Urteilsbildung im Geschichts- und Politikunterricht. Berlin 2010.
- KIENPOINTNER, MANFRED: Alltagslogik. Struktur und Funktion von Argumentationsmustern. Stuttgart-Bad Cannstatt 1992.
- KLAFKI, WOLFGANG: Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch konstruktive Didaktik. Weinheim/Basel ⁶2007.
- KLAFKI, WOLFGANG: Studien zur Bildungstheorie und Didaktik. Weinheim/Basel ^{37-40. Tsd.} 1975.
- KLEIN, E. DOMINIQUE/KRÜGER, MIRKO/KÜHN, SVENJA M./ACKEREN, ISABELL VAN: Wirkungen zentraler Abschlussprüfungen im Mehrebenensystem Schule. Eine Zwischenbilanz internationaler und nationaler Befunde und Forschungsdesiderate, in: Zeitschrift für Erziehungswissenschaft 17 (2014), 7–33.
- KLEIN, JOSEF: Erklären und Argumentieren als interaktive Gesprächskulturen, in: KLAUS BRINKER/GERD ANTOS/WOLFGANG HEINEMANN/SVEN F. SAGER (Hrsg.), Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung. Berlin/New York 2000, 1309–1329.
- KLIEME, ECKHARD/BRÜGELMANN, HANS: »Standards vorgeben?« Pro und Contra, in: Pädagogik 56 (2004), 50f.
- KLOSE, BRITTA: Diagnostische Wahrnehmungskompetenzen von ReligionslehrerInnen. Stuttgart 2014.
- KOHLBERG, LAWRENCE: Die Psychologie der Moralentwicklung. Hrsg. von WOLFGANG ALTHOF. Frankfurt a. M. 1996.
- KÖSTER, JULIANE: Leistung und Qualität von Korrekturanleitungen im Zentralabitur Deutsch, in: Der Deutschunterricht 1 (2010), 31–38.
- KÖSTER, JULIANE: Lern- und Leistungsaufgaben im Deutschunterricht, in: Deutschunterricht 61 (2008), 4–10.
- KUCKARTZ, UDO: Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. Weinheim/Basel ⁴2018.
- KÜHN, SVENJA M.: Aufgaben in (zentralen) Abschlussprüfungen. Theoretische und empirische Perspektiven auf ein interdisziplinäres Forschungsfeld, in: STEFAN KELLER/CHRISTIAN REINTJES (Hrsg.), Aufgaben als Schlüssel zur Kompetenz. Didaktische Herausforderungen, wissenschaftliche Zugänge und empirische Befunde. Münster 2016, 73–89.
- KÜHN, SVENJA M.: Steuerung und Innovation durch Abschlussprüfungen? Wiesbaden 2010.
- KUNIG, PHILIP: Rechtsfragen ethischer und religiöser Erziehung in der Schule, in: GOTTFRIED ADAM/FRIEDRICH SCHWEITZER (Hrsg.), Ethisch erziehen in der Schule. Göttingen 1996, 301–312.

- LACH, KURT: Planung von zwei Unterrichtsstunden: Der Castor-Transport – »Bürgerkrieg« im Wendland. War das Handeln der Beteiligten gerechtfertigt?, in: GOTTHARD BREIT (Hrsg.), Politische Urteilsbildung. Aufgabe und Wege für den Politikunterricht [= Schriftenreihe Bundeszentrale für politische Bildung, Bd. 344]. Bonn 1997.
- LENHARD, HARTMUT/PIRNER, MANFRED/ZIMMERMANN, MIRIAM: (Was) Hat der RU durch die Kompetenzorientierung gewonnen? Ein Schreibgespräch, in: Zeitschrift für Pädagogik und Theologie 70 (2018), 339–346.
- LIND, GEORG: Gewissen lernen? Zur Konstanzer Methode der Dilemmadiskussion, in: ANTON A. BUCHER (Hrsg.), Moral, Religion, Politik. Psychologisch-pädagogische Zugänge. Wien 2007, 101–111.
- LIND, GEORG: Inhalt und Struktur des moralischen Urteilens. Theoretische, methodologische und empirische Untersuchungen zur Moral- und Demokratiekompetenz bei Studierenden. Konstanz: 2000. Abrufbar unter: <http://kops.uni-konstanz.de/bitstream/handle/123456789/11305/Lind-Diss.pdf?sequence=1&isAllowed=y>.
- LINK, JÜRGEN: Versuch für den Normalismus. Wie Normalität produziert wird. Göttingen 1998.
- MASSING, PETER: Kategorien des politischen Urteilens und Wege zur politischen Urteilsbildung, in: DERS./GEORG WEIBENO (HRSG.), Politische Urteilsbildung, Schwalbach/Ts. 1997, 115–31.
- MATHES, STEPHANIE/KÜHN, SVENJA M.: »... dann kommt alles auf die Korrekturanleitung an...« Eine empirische Annäherung an ein (bislang) weitgehend vernachlässigtes Forschungsfeld, in: STEFAN KELLER/CHRISTIAN REINTJES (Hrsg.), Aufgaben als Schlüssel zur Kompetenz. Didaktische Herausforderungen, wissenschaftliche Zugänge und empirische Befunde. Münster 2016, 169–178.
- MAYRING, PHILIPP: Qualitative Inhaltsanalyse. Grundlagen und Techniken. Weinheim/Basel 2015.
- MEYER, CHRISTIANE/FELZMANN, DIRK: Was zeichnet ein gelungenes ethisches Urteil aus? Ethische Urteilskompetenz im Geographieunterricht unter der Lupe, in: CHRISTIANE MEYER/RODERICH HENRY/GEORG STÖBER (Hrsg.), Geographische Bildung. Kompetenzen in didaktischer Forschung und Schulpraxis. Braunschweig 2011, 130–146.
- MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT THÜRINGEN (Hrsg.), Gymnasiale Oberstufe in Thüringen. Voraussetzungen, Struktur, Unterrichtsfächer, Bewertung und Abiturprüfung. Erfurt: 2019. Abrufbar unter https://bildung.thueringen.de/fileadmin/ministerium/publikationen/gymnasiale_Oberstufe.pdf (13.5.2020).
- MINISTERIUM FÜR KULTUS, BILDUNG UND SPORT BADEN-WÜRTTEMBERG (Hrsg.), Bildungsplan des Gymnasiums. Geographie. Stuttgart 2016. Abrufbar unter: http://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lsbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_GYM_GEO.pdf (03.06.2020).
- MROCHEN, MARIA/HÖTTECKE, DIETMAR: Einstellungen und Vorstellungen von Lehrpersonen zum Kompetenzbereich Bewertung der Nationalen Bildungsstandards, in: Zeitschrift für interpretative Schul- und Unterrichtsforschung 1 (2012), 113–145.
- MÜLLER, MARGRIT: Die Förderung moralischer Urteilskompetenz im Biologieunterricht. Entwicklung eines Analyseinstruments zur Evaluation von Unterrichtskonzepten. Hamburg 2006.

- MUTH, KATHARINA: Orientierung geben und Identitäten stiften? Was der Religionsunterricht in Bezug auf Urteilsbildung leisten kann und sollte und was nicht, in: MIRKA DICKEL/ANKE JOHN/MICHAEL MAY/KATHARINA MUTH/LAURENZ VOLKMANN/MARIO ZIEGLER (Hrsg.), Urteilspraxis und Wertmaßstäbe im Unterricht. Ethik, Englisch, Geschichte, politische Bildung und Religion. Schwalbach 2020, 148–172.
- MUTH, KATHARINA: Über die Verortung und Bewertung von Einstellungen und Werthaltungen im Urteilsbildungsprozess. Zur religionspädagogischen Debatte mit einem Blick in die Sozialpsychologie, in: THOMAS HELLER (Hrsg.) Religion und Bildung – Interdisziplinär. Leipzig 2018, 515–526.
- NIEDERSÄCHSISCHES KULTUSMINISTERIUM (Hrsg.), Die gymnasiale Oberstufe und die Abiturprüfung. Informationen für Eltern und die Schülerinnen und Schüler, die ab 2021 ihr Abitur ablegen werden. Hannover 2019.
- NIEDERSÄCHSISCHES KULTUSMINISTERIUM (Hrsg.), Verordnung über die Abschlüsse in der gymnasialen Oberstufe, im Beruflichen Gymnasium, im Abendgymnasium und im Kolleg (AVO-GOBAG) vom 19. Mai 2005 und geändert durch Verordnung vom 04. September 2018.
- NIEDERSÄCHSISCHES KULTURMINISTERIUM (Hrsg.), Kerncurriculum für das Gymnasium – gymnasiale Oberstufe, die Gesamtschule – gymnasiale Oberstufe, das Berufliche Gymnasium, das Kolleg. Evangelische Religion. Hannover 2017.
- NIPKOW, KARL E.: Bildung in einer Pluralen Welt. Bd. II: Religionspädagogik im Pluralismus. Gütersloh 1998.
- NIPKOW, KARL E.: Religionsunterricht in der Leistungsschule. Gutachten, Dokumente. Gütersloh 1979.
- OBST, GABRIELE: Kompetenzorientiertes Lehren und Lernen im Religionsunterricht. Göttingen ⁴2015.
- ODENDAHL, JOHANNES: Zentral, nicht zerebral. Wider die Benotung zentral gestellter Klausuren anhand eines Bewertungsrasters, in: Pädagogische Rundschau 62 (2008), 451–457.
- OSER, FRITZ/BIEDERMANN, HORST: Art. Moralentwicklung und Moralförderung, in: Handwörterbuch Pädagogische Psychologie 2018, 559–568.
- OTTO, STEPHAN: »Interpretieren Sie das Gedicht...« Aufgaben in zentralen Abschlussprüfungen im Fach Deutsch, in: STEFAN KELLER/CHRISTIAN REINTJES (Hrsg.), Aufgaben als Schlüssel zur Kompetenz. Didaktische Herausforderungen, wissenschaftliche Zugänge und empirische Befunde. Münster/New York 2016, 142–151.
- OTTO, STEPHAN/ KÜHN, SVENJA M.: Zwischen Tradition und Innovation. Eine Analyse zentraler Prüfungsaufgaben im Fach Deutsch im Nationalen Vergleich, in: Beziehungen in Unterricht und Schule 5 (2014), 1–16.
- PIRKER, VIERA/JUEN, MARIA: Religion – (k)ein Fach wie jedes andere. Spannungsfelder und Perspektiven in der kompetenzorientierten Leistungsbeurteilung. Stuttgart 2018.
- PIRNER, MANFRED L.: Religiosität und Lehrerprofessionalität. Ein Literaturbericht zu einem vernachlässigten Forschungsfeld, in: Zeitschrift für Pädagogik 59 (2013), 201–218.
- PLATZER, KATRIN/SINEMUS, KRISTINA: Ein Strukturmodell ethischer Urteilsbildung im Kontext moderner Biotechnologie und Gentechnik. Das Fallbeispiel Bt-Mais, in: CHRISTINE HAUSKELLER (Hrsg.), Wissenschaft verantworten. Soziale und ethische Orientierung in der technischen Zivilisation. Münster 2001, 91–107.

- POHL-PATALONG, UTA (Hrsg.), Kompetenzorientierung im Religionsunterricht – Chancen und Grenzen, in: Zeitschrift für Pädagogik und Theologie 70 (2018), 372–385.
- POLLACK, DETLEF/MÜLLER, OLAF: Religionsmonitor. Verstehen was verbindet. Religiosität und Zusammenhalt in Deutschland. Gütersloh 2013.
- PREIBER, RÜDIGER: Kompetenzorientierte Hochschuldidaktik, in: FLORIAN BRUCKMANN/OLIVER REIS/MONIKA SCHEIDLER (Hrsg.), Kompetenzorientierte Lehre in der Theologie. Konkretion – Reflexion – Perspektiven. Berlin 2011, 17–36.
- PREIBER, RÜDIGER/VÖLZKE, REINHARD: Kompetenzbilanzierung – Hintergründe, Verfahren, Entwicklungsnotwendigkeit, in: REPORT 30 (2007), 62–71. Abrufbar unter: <https://www.die-bonn.de/doks/preisser0701.pdf> (14. 6.2020).
- PRIEMER, BURKHARD: Deutschsprachige Verfahren der Erfassung von epistemologischen Überzeugungen, in: Zeitschrift für Didaktik der Naturwissenschaften 12 (2006), 159–175.
- REIS, OLIVER: Zwischen studentischer Diversität und fachlicher Homogenisierung. Die wertbezogene Hochschuldidaktik als Vermittlerin im Habituskonflikt, in: SYLVIA HEUCHEMER/FRIEDERIKE SILLER/TIMO VAN TREECK (Hrsg.), Hochschuldidaktik forscht zu Vielfalt und Offenheit. Profilbildung und Wertefragen in der Hochschulentwicklung I. Cologne Open Science: 2018, 19–49. Abrufbar unter: https://cos.bibl.thkoeln.de/frontdoor/deliver/index/docId/802/file/FIHB_Band_2_final_web.pdf (23. 1. 2019).
- REIS, OLIVER: Sinn und Umsetzung der Kompetenzorientierung – Lehre »von hinten« denken, in: PATRICK BECKER (Hrsg.), Studienreform in der Theologie. Eine Bestandsaufnahme. Berlin 2011, 108–127.
- REIS, OLIVER/SCHWARZKOPF, THERESA: Diagnosemodelle für den Religionsunterricht. Eine Bestandsaufnahme, in: Religionspädagogische Beiträge 76 (2017), 85–95.
- REIS, OLIVER/SCHWARZKOPF, THERESA: Diagnose religiöser Lernprozesse. Ein kompetenzdiagnostisches Grundlagenmodell, in: DIES. (Hrsg.), Diagnose im Religionsunterricht. Konzeptionelle Grundlagen und Praxiserprobungen. Berlin 2015, 15–120.
- REITSCHERT, KATJA/LANGLÉT, JÜRGEN/HÖBLE, CORINNA/ MITTELSTEN SCHEID, NICOLA/SCHLÜTER, KIRSTEN: Dimensionen Ethischer Urteilskompetenz. Dimensionierung und Niveaunkonkretisierung, in: Der mathematische und naturwissenschaftliche Unterricht 60 (2007), 43–51.
- REUTER, HANS-RICHARD: Grundlagen und Methodik der Ethik, in: WOLFGANG HUBER/TORSTEN MEIREIS/HANS-RICHARD REUTER (Hrsg.), Handbuch der Evangelischen Ethik. München 2015.
- ROTH, MICHAEL: Ist Moral lehrbar? Überlegungen zur ethischen Kompetenz, in: BERND HARBECK-PINGEL/MICHAEL ROTH (Hrsg.), Emotionen und Verhalten in theologischer und philosophischer Perspektive. Leipzig 2012, 189–202.
- ROTHGANGEL, MARTIN: Das Beste am Religionsunterricht? Kompetenzorientierung im anthropologischen Horizont, in: THOMAS SCHLAG/HENRIK SIMOJOKI (Hrsg.), Mensch – Religion – Bildung. Religionspädagogik in anthropologischen Spannungsfeldern. Gütersloh 2014, 478–489.
- ROTHGANGEL, MARTIN: Bildungsstandards Religion. Eine Replik auf verbreitete Kritikpunkte, in: ANDREAS FEINDT/VOLKER ELSENBAST/PETER SCHREINER/ALBRECHT SCHÖLL

- (Hrsg.), Kompetenzorientierung im Religionsunterricht. Befunde und Perspektiven. Münster u. a. 2009, 87–97.
- ROTHGANGEL, MARTIN/FISCHER, DIETLIND (Hrsg.), Standards für religiöse Bildung? Zur Reformdiskussion in Schule und Bildung. Münster 2004.
- SACHER, WERNER: Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe. Bad Heilbrunn ⁵2009, 36–42.
- SAJAK, CLAUB PETER: Religionsunterricht kompetenzorientiert: Beiträge aus fachdidaktischer Forschung. Paderborn u. a. 2012.
- SALDERN, MATTHIAS VON: Schulleistung 2.0. Von der Note zum Kompetenzraster. Norderstedt 2011.
- SCHAPER, NICLAS: Ansätze zur Kompetenzmodellierung und -messung im Rahmen einer kompetenzorientierten Hochschuldidaktik, in: FLORIAN BRUCKMANN/OLIVER REIS/MONIKA SCHEIDLER (Hrsg.), Kompetenzorientierte Lehre in der Theologie. Konkretion – Reflexion – Perspektiven. Berlin 2011, 37–62.
- SCHMIDT, CHRISTIANE: Analyse von Leitfadenterviews, in: BARBARA FRIEBERTSHÄUSER/ANTJA LANGER/ANNEDORE PRENGEL (Hrsg.), Handbuch. Qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim ⁴2013, 473–486.
- SCHOBERTH, INGRID: Urteilen lernen. Einleitende Reflexionen, Perspektiven und Orientierungen in religionspädagogischer Perspektive, in: DIES. (Hrsg.), Urteilen lernen – Grundlegung und Kontexte ethischer Urteilsbildung. Göttingen/Oakville 2012, 25–40.
- SCHÖNEMANN, BERND/THÜNEMANN, HOLGER/ZÜLSDORF-KERSTING, MEIK: Was können Abiturienten? Zugleich ein Beitrag zur Debatte über Kompetenzen und Standards im Fach Geschichte. Berlin/Münster ²2011.
- SCHRÖDER, BERND: Religion unterrichten in Niedersachsen, in: MARTIN ROTHGANGEL/BERND SCHRÖDER (Hrsg.), Religionsunterricht in den Ländern der Bundesrepublik Deutschland. Neue empirische Daten – Kontexte – Aktuelle Entwicklungen. Leipzig 2020, 239–267.
- SCHRÖDER, BERND: Was macht es für einen Unterschied, ob ethische Fragen im Ethik- oder evangelischen / katholischen / jüdischen / islamischen Religionsunterricht behandelt werden? In: RUDOLF ENGLERT/HELGA KOHLER-SPIEGEL/ELISABETH NAURATH/BERND SCHRÖDER/FRIEDRICH SCHWEITZER (Hrsg.), Ethisches Lernen. Neukirchen-Vluyn: 2015 [= Jahrbuch der Religionspädagogik 31], 41–63.
- SCHWAIHOFER, MATTHIAS /HEENE, MORITZ/BÜHNER, MARKUS: Grundlagen und Kriterien der Diagnostik, in: DETLEF URHAHNE/MARKUS DRESSEL/FRANK FISCHER (Hrsg.), Psychologie für den Lehrberuf. Berlin/Heidelberg 2019, 493–516.
- SCHWARZKOPF, THERESA: Vielfältigkeit Denken. Wie Schülerinnen und Schüler im Religionsunterricht argumentieren lernen. Stuttgart 2016.
- SCHWEITZER, FRIEDRICH: Leistungsmessung und Leistungsbewertung, Kompetenzen und Standards: Was ist im Religionsunterricht messbar? In: entwurf 47 (2016), 4–8.
- SCHWEITZER, FRIEDRICH: Religiöse Bildung ohne Ethik? Zur ethischen Dimension des Religionsunterrichts, in: RUDOLF ENGLERT/HELGA KOHLER-SPIEGEL/ELISABETH NAURATH/BERND SCHRÖDER/FRIEDRICH SCHWEITZER (Hrsg.), Ethisches Lernen. Neukirchen-Vluyn 2015 [= Jahrbuch der Religionspädagogik 31], 13–23.

- SPINNER, KASPAR: Der standardisierte Schüler, in: Informationen zur Deutschdidaktik 31 (2007), 130–133.
- STATISTISCHES BUNDESAMT (Hrsg.), Pressemitteilung. Schülerzahlen im Schuljahr 2017/18 um 0,4 % zurückgegangen. Wiesbaden 2018. Abrufbar unter: https://www.destatis.de/DE/Presse/Pressemitteilungen/2018/03/PD18_089_211.html (26.3.2020).
- STAATSMINISTERIUM FÜR SCHULQUALITÄT UND BILDUNGSFORSCHUNG MÜNCHEN (Hrsg.), Lehrplan. Evangelische Religionslehre. Jahrgang 11/12. München 2003. Abrufbar unter: <http://www.isb-gym8lehrplan.de/contentserv/3.1.neu/g8.de/index.php?StoryID=26176> (02.09.2020).
- STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BUNDESREPUBLIK DEUTSCHLAND IVC/STATISTIK (Hrsg.), Abiturnoten 2019 an Gymnasien, Integrierten Gesamtschulen, Fachgymnasien, Fachoberschulen und Berufsschulen. Berlin 2020. Abrufbar unter: <https://www.kmk.org/dokumentation-statistik/statistik/schulstatistik/abiturnoten.html> (07.2.2020).
- STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BUNDESREPUBLIK DEUTSCHLAND IVC/STATISTIK (Hrsg.), Auswertung Religionsunterricht Schuljahr 2017/18. Teilnehmende Schülerinnen und Schüler allgemeinbildender Schulen in öffentlicher Trägerschaft nach Schularten (aufgegliedert nach Religionsunterricht, Ethik und weiteren Ersatzunterricht) für den Primar- und Sekundarbereich I. Berlin 2019. Abrufbar unter: https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/AW_Religionsunterricht_II_2017_18.pdf (26.3.2019).
- STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER (Hrsg.), Einheitliche Anforderungen für die Abiturprüfung im Fach Geschichte. Beschluss vom 1.12.1989 in der Fassung vom 10.2.2015.
- STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER (Hrsg.), Vereinbarung über Einheitliche Prüfungsanforderungen in der Abiturprüfung. Beschluss der Kultusministerkonferenz vom 01.06.1979 in der Fassung vom 24.10.2008.
- STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER (Hrsg.), Einheitliche Prüfungsanforderungen in der Abiturprüfung Evangelische Religionslehre. Beschluss vom 01.12.1989 i. d. F. vom 16.11.2006. Abrufbar unter: https://www.kmk.org/fileadmin/veroeffentlichungen_beschlue-sse/1989/1989_12_01-EPA-Ev-Religion.pdf (13.04.2019).
- STÄNDIGE KONFERENZ DER KULTUSMINISTER DER LÄNDER (Hrsg.), Beschlüsse der Kultusministerkonferenz. Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004. München/Neuwied 2005.
- STEIN, ANDREAS: »Zentralabitur« in Niedersachsen. Zur Entstehung von Prüfungsaufgaben mit landesweit einheitlichen Aufgabenstellungen, in: SchulVerwaltung Niedersachsen 1 (2020), 22–24.
- STEINHÄUSER, MARTIN: Art. Christenlehre, in: Wissenschaftliches Religionspädagogisches Lexikon im Internet (www.wirelex.de) 2016.
- THEIN, CHRISTIAN: Problemreflexion und Urteilsbildung im Philosophieunterricht, in: Mitteilungen des Fachverbandes Philosophie e.V. 56 (2016), 85–93. Abrufbar unter: http://fv-philosophie-nrw.de/wp-content/uploads/2016/12/Mitteilungen_gesamt_2016_web.pdf (03.07.2019).

- THÜRINGER MINISTERIUM FÜR BILDUNG, WISSENSCHAFT UND KULTUR (Hrsg.), Lehrplan für den Erwerb der allgemeinen Hochschulreife. Evangelische Religionslehre. 2013.
- THÜRINGER MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT (Hrsg.), Thüringer Schulordnung. Thüringer Schulordnung für die Grundschule, die Regelschule, die Gemeinschaftsschule, das Gymnasium und die Gesamtschule vom 20. Januar 1994.
- TILICH, PAUL: Gesammelte Werke. Bd. 9, Stuttgart 1967.
- TÖDT, HEINZ EDUARD: Perspektiven theologischer Ethik. München 1988.
- TOULMIN, STEPHEN: Der Gebrauch von Argumenten. Weinheim ²1996.
- ULRICH-RIEDHAMMER, EVA-MARIE/APPLIS, STEFAN: Ethisches Argumentieren als Herausforderung, in: Praxis Geographie 3 (2013) 24–29.
- URHAHNE, DETLEF/DRESSSEL, MARKUS/FISCHER, FRANK (Hrsg.), Psychologie für den Lehrberuf. Berlin/Heidelberg 2019.
- VISSER, ELKE: Die Diagnose der Bewertungskompetenz durch schriftliche Aufgaben im Biologieunterricht. Hamburg 2014.
- VISSER, ELKE/HÖBLE, CORINNA: Bioethisch argumentieren – Ein Diagnostischer Blick auf die Bewertungskompetenz im Biologieunterricht, in: ALEXANDRA BUDKE/MIRIAM KUCKUCK/MICHAEL MAYER/FRANK SCHÄBITZ/KIRSTEN SCHLÜTER/GÜNTHER WEISS (Hrsg.), Fachlich argumentieren lernen. Didaktische Forschungen zur Argumentation in den Unterrichtsfächern. Münster/New York 2015, 182–195.
- WALDEGGE, KERSTIN H. VON: Biologielehrkräfte diagnostizieren die Schülerkompetenz »Bewerten«. Eine qualitative Untersuchung zur Orientierung bei der Diagnose. Hamburg 2016.
- WEHLING, HANS-GEORG: Konsens á la Beutelsbach? Nachlese zu einem Expertengespräch, in: SIEGFRIED SCHIELE/HERBERT SCHNEIDER (Hrsg.), Das Konsensproblem in der politischen Bildung. Stuttgart 1977, 173–184.
- WEINBRENNER, PETER: Politische Urteilsbildung als Ziel und Inhalt des Politikunterrichts, in: BUNDESZENTRALE FÜR POLITISCHE BILDUNG (Hrsg.), Politische Urteilsbildung. Aufgaben und Wege für den Politikunterricht. Bonn 1997, 73–94.
- WEINERT, FRANZ E.: Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit, in: DERS. (Hrsg.), Leistungsmessungen in Schulen. Weinheim 2001, 17–31.
- WEIß, THOMAS: Fachspezifische und fachübergreifende Argumentationen am Beispiel von Schöpfung und Evolution. Theoretische Grundlagen – empirische Analysen – Jugendtheologische Konsequenzen. Göttingen 2016.
- WEIß, THOMAS/BASEL, NICOLAI/HARMS, UTE/PRECHTL, HELMUT/ROTHGANGEL, MARTIN, Fächerübergreifende Kompetenzen im Unterricht. Argumentieren in den Fächern Biologie und Religion – ein Beispiel, in: UTE HARMS/BURKHARD SCHROETER/BARBARA KLÜH (Hrsg.), Entwicklung kompetenzorientierten Unterrichts in Zusammenarbeit von Forschung. Münster 2016, 205–223.
- WERMKE, MICHAEL (Hrsg.), Bildungsstandards und Religionsunterricht. Perspektiven aus Thüringen. Jena 2005.
- WERMKE, MICHAEL/WIDL, MARIA: Religion unterrichten in Thüringen, in: MARTIN ROTHGANGEL, BERND SCHRÖDER (Hrsg.), Religionsunterricht in den Ländern der Bundesrepublik Deutschland. Neue empirische Daten – Kontexte – Aktuelle Entwicklungen. Leipzig 2020, 419–442.

- WINKLER, IRIS: Argumentierendes Schreiben im Deutschunterricht im Spiegel von Aufgaben für Lern- und Leistungssituationen, in: ELKE GRUNDLER/RÜDIGER VOGT (Hrsg.), Argumentieren in Schule und Hochschule. Interdisziplinäre Studien. Tübingen 2006, 157–166.
- WORBACH, MARC/DRECHSEL, BARBARA/CARSTENSEN, CLAUS H.: Messen und Bewerten von Lernergebnissen, in: DETLEF URHAHNE/MARKUS DRESSSEL/FRANK FISCHER (Hrsg.), Psychologie für den Lehrberuf. Berlin/Heidelberg 2019, 493–516.
- ZABKA, THOMAS/STARK, TOBIAS: Aufgabenstellungen und Erwartungshorizonte als Steuerungsinstrumente. Zum Umgang mit Problemen der Literaturinterpretation im Zentralabitur, in: Der Deutschunterricht 1 (2010), 19–29.
- ZIEGLER, MARIO: Die Schulung der Urteilsfähigkeit im Ethikunterricht. Grundsätze der Lehrstückdidaktik, in: MIRKA DICKEL/ANKE JOHN/MICHAEL MAY/KATHARINA MUTH/LAURENZ VOLKMANN/MARIO ZIEGLER (Hrsg.), Urteilspraxis und Wertmaßstäbe im Unterricht. Ethik, Englisch, Geschichte, politische Bildung und Religion. Schwalbach 2020, 47–70.
- ZIEGLER, MARIO: Die festgefügtten Urteile des moralisch Naiven. Möglichkeiten der Reflexion und Revision am Beispiel von Kurosawas Film »Rashomon«, in: MARKUS TIEDEMANN (Hrsg.), Werte und Wertevermittlung. Dresden 2018, 63–73.
- ZWERGEL, HERBERT A.: Kompetenzen und Grenzen ihrer Standardisierung – Umgang mit ethischen Entscheidungssituationen und existenziellen Grenzsituationen am Beispiel religiösen Lernens, in: VÖLKER FREDERKING (Hrsg.), Schwer messbare Kompetenzen. Herausforderungen für die empirische Fachdidaktik. Baltmannsweiler 2008, 85–114.

Abbildungsverzeichnis

- Abb. 1: Fragestellungen ethischer und religiöser Urteilsbildung.
- Abb. 2: Unterscheidung zwischen Sach- und Werturteil.
- Abb. 3: Schematische Darstellung eines kompetenzorientierten Bildungsverständnisses.
- Abb. 4: Die klassischen Gütekriterien der Testtheorie.
- Abb. 5: Vertikale Lernstandsdiagnostik und horizontale Lernprozessdiagnostik. Nachbildung aus REIS/SCHWARZKOPF, Diagnose religiöser Lernprozesse, 77.
- Abb. 6: Hermeneutischer Kreisprozess zur Beurteilung ganzheitlicher Leistungen. Eigene Darstellung nach SACHER: Leistungen entwickeln, überprüfen und bewerten, 155.
- Abb. 7: Das dreiteilige Modell einer Argumentation.
- Abb. 8: Die der KOHLBERG'schen Stufentheorie zugrunde liegende Gewichtung von Argumenten am Beispiel der Fragestellung, ob Diebstahl in prekären Situationen legitim sei (angelehnt an das sog. »Heinz-Dilemma«).
- Abb. 9: Bearbeitungsstruktur zur Argumentationskompetenz in metaphysisch-religiösen Fragen. Nachbildung aus SCHWARZKOPF, Vielfältigkeit denken, 179.
- Abb. 10: »Ablaufschema einer inhaltlich strukturierenden Inhaltsanalyse«. Übernommen aus: KUCKARTZ, Qualitative Inhaltsanalyse, 100.
- Abb. 11: Säulendiagramm zur durchschnittlichen Anzahl der Aufgabenstellungen je Aufgabenset (2014–2019) im schriftlichen Abitur im Fach Evangelische Religionslehre unter Angabe der Bearbeitungszeit.
- Abb. 12: Zusammenfassendes Kreisdiagramm zu den Aufgabenarten urteilsbildender Aufgabenstellungen von 2014–2019 in den Bundesländern Bayern, Niedersachsen und Thüringen.
- Abb. 13: Kreisdiagramm über die Fragestellungen der Urteilsbildung in Bayern.
- Abb. 14: Kreisdiagramm über die Fragestellungen der Urteilsbildung in Niedersachsen.

- Abb. 15: Kreisdiagramm über die Fragestellungen der Urteilsbildung in Thüringen.
- Abb. 16: Reflexionskriterien der Urteilsbildung im Untersuchungszeitraum 2014–2019.
- Abb. 17: Säulendiagramm zur relativen Häufigkeit von Urteilsbildungen im schriftlichen Abitur in den Bundesländern Bayern, Niedersachsen und Thüringen.
- Abb. 18: Qualitätsmerkmale für Urteilsbildung, die in mindestens zwei der untersuchten Bundesländer mehr als einmal genannt werden.
- Abb. 19: Charakteristika der Aufgabenkultur und Bewertungsvorgaben von Urteilsbildungsaufgaben aus Bayern.
- Abb. 20: Charakteristika der Aufgabenkultur und Bewertungsvorgaben von Urteilsbildungsaufgaben aus Niedersachsen.
- Abb. 21: Charakteristika der Aufgabenkultur und Bewertungsvorgaben von Urteilsbildungsaufgaben aus Thüringen.

Tabellenverzeichnis

- Tab. 1: Für Urteilsfähigkeit relevante Operatoren nach den EPA, 14 f.
- Tab. 2: Niveaustufenmodell moralischer Kompetenz nach GEORG LIND. Eigene Tabelle nach LIND, Gewissen lernen?, 101–111.
- Tab. 3: Die im Projekt ETiK identifizierten Niveaustufen ethisch-moralischer Urteilskompetenz. Eigene Tabelle nach BENNER/NIKOLOVA, Anforderungsniveaus der ethisch-moralischen Teilkompetenzen, 141–144.
- Tab. 4: Niveaustufen für die Teilkompetenz Argumentieren. Eigene Tabelle nach VISSER, Diagnose der Bewertungskompetenz, 280.
- Tab. 5: Niveaustufen der deskriptiven Aspekte. Eigene Tabelle nach VISSER, Diagnose der Bewertungskompetenz, 289.
- Tab. 6: Niveaustufen normative Aspekte. Eigene Tabelle nach VISSER, Diagnose der Bewertungskompetenz, 278.
- Tab. 7: Empirisch überprüfte Qualitätsmerkmale ethischer und religiöser Urteilskompetenz im Überblick.
- Tab. 8: Gesamtdurchschnitt der Abiturnoten in den Bundesländern Bayern, Niedersachsen und Thüringen. Eigene Tabelle auf Grundlage des Sekretariats der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland IVC/Statistik.
- Tab. 9: Vergleich der Teilnehmerinnen und Teilnehmerzahlen am Religionsunterricht in Bayern, Niedersachsen und Thüringen im Schuljahr 2017/18.
- Tab. 10: Anzahl der Prüflinge schriftlicher Abiturprüfungen in den Fächern Evangelische Religionslehre und Ethik im Bundesland Thüringen.
- Tab. 11: Anzahl der Prüflinge schriftlicher Abiturprüfungen in den Fächern Evangelische Religionslehre und Ethik im Bundesland Bayern.
- Tab. 12: Anzahl der Prüflinge schriftlicher Abiturprüfungen in den Fächern Evangelische Religionslehre und Werte und Normen im Bundesland Niedersachsen.

- Tab. 13: Kategoriensystem zur Urteilsfähigkeit in kontroversen ethischen und religiösen Fragen im schriftlichen Abitur des Faches Evangelische Religionslehre.
- Tab. 14: Summe der Aufgabenstellungen im schriftlichen Abitur des Faches Evangelische Religionslehre in den Jahren 2014–2019.
- Tab. 15: Summe und relative Häufigkeit der Aufgabenstellungen zu ethischer und/oder religiöser Urteilskompetenz im schriftlichen Abitur des Faches Evangelische Religionslehre in den Jahren 2014–2019.
- Tab. 16: Summe und relative Häufigkeit der Aufgabensets, die mindestens eine Aufgabenstellung zu ethischer und/oder religiöser Urteilskompetenz im schriftlichen Abitur des Faches Evangelische Religionslehre in den Jahren 2014–2019 aufweisen.
- Tab. 17: Systematisierung der Fragestellungen der Urteilsbildung.
- Tab. 18: Fragestellungen der Urteilsbildung in Bayern im Zeitraum von 2014–2019.
- Tab. 19: Fragestellungen der Urteilsbildung in Niedersachsen im Zeitraum von 2014–2019.
- Tab. 20: Fragestellungen der Urteilsbildung in Thüringen im Zeitraum von 2014–2019.
- Tab. 21: Qualitätsmerkmale der Urteilsbildung aus den Erwartungshorizonten bayerischer Abituraufgabenstellungen in den Jahren 2014–2019.
- Tab. 22: Qualitätsmerkmale der Urteilsbildung aus den Erwartungshorizonten niedersächsischer Abituraufgabenstellungen in den Jahren 2014–2019.
- Tab. 23: Qualitätsmerkmale der Urteilsbildung aus den Erwartungshorizonten thüringischer Abituraufgabenstellungen in den Jahren 2014–2019.

Abkürzungsverzeichnis

a. a. O.	am angegebenen Ort
Abb.	Abbildung
AFB	Anforderungsbereiche
Anm.	Anmerkung
BE	Bewertungseinheiten
BMBF	Bundesministerium für Bildung und Forschung
bspw.	beispielsweise
BY	Bayern
Dass.	Dasselbe
Ders.	Derselbe
Dies.	Dieselbe/n
eA	erhöhtes Anforderungsniveau
EH	Erwartungshorizont
EKD	Evangelische Kirche in Deutschland
EPA	Einheitliche Prüfungsanforderungen
ebd.	ebenda
etc.	et cetera
ev.	evangelisch
gA	grundlegendes Anforderungsniveau
ggf.	gegebenenfalls
Hrsg.	Herausgeber/Herausgeberin
Kap.	Kapitel
LP	Lehrplan
NI	Niedersachsen
sic	<i>sic erat scriptum</i> – so stand es geschrieben
Tab.	Tabelle
TH	Thüringen
u. a.	unter anderem
UB	Urteilsbildung
vgl.	vergleiche
z. B.	zum Beispiel

Anhang

Anhang I: Berücksichtigte Operatoren aus den EPA¹

Operatoren	Definition
Sich auseinandersetzen mit	ein <i>begründetes eigenes Urteil</i> zu einer Position oder einem dargestellten Sachverhalt entwickeln
Beurteilen Bewerten Stellung nehmen Einen begründeten Standpunkt einnehmen	zu einem Sachverhalt unter Verwendung von Fachwissen und Fachmethoden sich <i>begründet positionieren</i> (Sach- bzw. Werturteil)
Erörtern	die Vielschichtigkeit eines Beurteilungsproblems erkennen und darstellen, dazu Thesen erfassen bzw. aufstellen, Argumente formulieren, nachvollziehbare Zusammenhänge herstellen und dabei eine <i>begründete Schlussfolgerung erarbeiten</i> (dialektische Erörterung)
Prüfen Überprüfen	eine Meinung, Aussage, These, Argumentation nachvollziehen, kritisch befragen und auf der Grundlage erworbener Fachkenntnisse <i>begründet beurteilen</i>
Gestalten Entwerfen	sich textbezogen mit einer Fragestellung kreativ auseinander setzen
Stellung nehmen aus der Sicht von... Eine Erwiderung formulieren aus der Sicht von...	eine unbekannte Position, Argumentation oder Theorie aus der Perspektive einer bekannten Position beleuchten oder in Frage stellen und ein <i>begründetes Urteil</i> abgeben
Konsequenzen aufzeigen Perspektiven entwickeln	Schlussfolgerungen ziehen; Perspektiven, Modelle, Handlungsmöglichkeiten, Konzepte u. a. entfalten

¹ EPA, 14 f., Hervorhebungen K.M.

Anhang II: Die von den EPA vorgegebenen Bewertungskriterien

- »Umfang der Kenntnisse und Einsichten
- Breite der Argumentation
- Vielfalt der Aspekte
- Reichhaltigkeit der Argumente
- Qualität der Beispiele
- Kreativität im Umgang mit Problemstellungen und Gestaltungsaufgaben

- Sachgemäßheit und Komplexität des Text- bzw. Problemverständnisses
- Genauigkeit der Kenntnisse
- Stimmigkeit der Darstellung
- Klarheit der Gedankenführung

- Gliederung der Darstellung
- Differenzierung zwischen Wichtigem und weniger Wichtigem
- Reflexionsniveau

- Beherrschung von Fachmethoden
- sachgemäßer Umgang mit Fachsprache
- Klarheit des Ausdrucks
- begriffliche Exaktheit
- sprachliche Angemessenheit.«²

² EPA., 20.

Anhang III: Die Fälle der Studie (42) – Aufgabenstellungen zu Urteilsfähigkeit in kontroversen ethischen oder religiösen Fragestellungen aus der Auswahlinheit

Bez.	Aufgabenstellung	Signatur
B1	In Ihrem Ort wird ein neues behindertengerechtes Sportzentrum eröffnet. Entwerfen Sie einen Beitrag für die Vereinszeitschrift des Sportvereins mit dem Titel »Schneller – besser – höher: aber natürlich!«, in dem Sie differenziert auf die Frage der Selbstoptimierung im Sport eingehen! Beziehen Sie dabei in kritischer Lektüre den Text (Seite 4f.) sowie Gedanken eines christlichen Menschenbildes ein!	[BY, 2014 (2)]
B2	4 An Ihrer Schule wird die Ausstellung »Local heroes« gezeigt. Sie sind als Schülersvertretung damit beauftragt, eine Rede bei der Eröffnung der Ausstellung zu halten. Entwerfen Sie eine Rede mit dem Titel »Tolle Typen, wozu? – Brauchen wir wirklich Vorbilder?« Ziehen Sie dazu die Materialien 1–4 heran und berücksichtigen Sie dabei evangelische Einsichten zu Gott und Mensch!	[BY, 2014 (4)]
B3	3. An Ihrer Schule überarbeitet ein Team aus Lehrkräften, Eltern, Schülerinnen und Schülern die Schulverfassung. In der letzten Sitzung dieser Gruppe wurde kontrovers diskutiert, ob in den für die Schule gültigen Wertekanon auch »Demut« aufgenommen werden soll. Eine Entscheidung darüber steht noch aus. Entwerfen Sie eine E-Mail an die Arbeitsgruppe, in der Sie Ihren Standpunkt darlegen! Berücksichtigen Sie dabei die Materialien 1–4 [Zeitungsartikel und eine Leserbrief – Meinungen]! [30 BE]	[BY, 2015 (3)]
B4	5. Dankbarkeit und Demut (vgl. Material 1, Z. 18–20) auch im Leid? Entfalten Sie dazu eine Antwort und greifen Sie dabei auf einen Ansatz zur Theodizee zurück! [15 BE]	[BY, 2015 (3)]
B5	2.2 Nehmen Sie zu einer von Grafts zentralen Aussagen kritisch Stellung! [10 BE]	[BY, 2015 (4)]
B6	3 Beurteilen Sie die Position Hubers zum Einsatz von Gewalt! Berücksichtigen Sie dabei Grundzüge biblischer Ethik sowie Denkmodelle zum politischen Engagement von Christen! [25 BE]	[BY, 2016 (4)]

Fortsetzung

Bez.	Aufgabenstellung	Signatur
B7	4 »Die Achtung vor der Wahrheit, die Wahrung der Menschenwürde und die wahrhaftige Unterrichtung der Öffentlichkeit sind oberste Gebote der Presse.« (Ziffer 1, Pressekodex des Deutschen Presserates) Erörtern Sie Möglichkeiten und Grenzen der Umsetzung dieses Grundsatzes unter Berücksichtigung der im Text (Seite 2 f.) entfalteten Gedanken! Beziehen Sie dabei grundlegende Aspekte philosophischer und christlicher Ethik mit ein! [25 BE]	[BY, 2017 (1)]
B8	4 »Wir müssen lernen, damit umzugehen. Mit unserem eigenen Neid – und dem der anderen.« (Material 1, Zeile 57 f.) Entwerfen Sie einen Artikel für die Schülerzeitung über den Umgang mit Neid im Schulleben! Beziehen Sie Gedanken der Materialien 1–4 und Aspekte biblischer Anthropologie und Ethik mit ein! [30 BE]	[BY, 2017 (3)]
B9	3 Sie haben an einem W-Seminar mit Leitfach »Evangelische Religionslehre« zum Thema »Zukunfts-Visionen« teilgenommen, in dem unterschiedliche (biologische, politische, religiöse usw.) utopische Vorstellungen untersucht und bewertet wurden.³ Ausgewählte Ergebnisse werden bei einem Vortragsabend für Eltern und interessiertes Publikum vorgestellt. Entwerfen Sie einen Redebeitrag, der sich mit dem Konzept der Kryonik (vgl. Text) auseinandersetzt! Berücksichtigen Sie dabei Aspekte christlichen Menschenverständnisses (Aufgabe 2) sowie christlicher Zukunftshoffnung! [30 BE]	[BY, 2017 (4)]
B10	Die Redaktion der Schülerzeitung Ihrer Schule plant eine Thementausgabe mit dem Titel »Leben in der Welt von morgen«. Dabei sollen die unterschiedlichen Fächer zu Wort kommen. Entwerfen Sie für das Fach Evangelische Religionslehre einen Beitrag, der sich mit Chancen und Grenzen des technisch Machbaren auseinandersetzt! Gehen Sie dabei von zentralen Gedanken des Textes auf Seite 12 f. aus, wählen Sie aber eine andere Thematik als die des autonomen Fahrens! Formulieren Sie für Ihren Textbeitrag eine passende Überschrift! [30]	[BY, 2018 (4)]

³ Ein »W-Seminar« ist eines von zwei Seminaren, welches Schüler und Schülerinnen in Bayern in der Oberstufe belegen. Es dient zur Vorbereitung auf wissenschaftliches Arbeiten. Darüber hinaus gibt es P-Seminare, welche sich mit der Berufsorientierung befassen.

Fortsetzung

Bez.	Aufgabenstellung	Signatur
B11	5 Setzen Sie sich damit auseinander, inwieweit grundlegende Aussagen des Apostolischen Glaubensbekenntnisses »die letzten Fragen nach dem Sinn unserer Existenz, dem Warum, Woher, Wozu« (Z. 40 f.) umfassend beantworten können.	[BY, 2019 (4)]
T1	1. Entwerfen Sie in Auseinandersetzung mit den Worten von Margot Käßmann eine Rede zur Abschlussfeier der Absolventen Ihres Jahrgangs. [Aussage zur politischen Dimension von Geboten und biblischer Texte und Bekenntnis zur »Weltverbesserin«] Beziehen Sie die Materialien, geeignete Texte aus der Bibel sowie Ihre persönlichen Erfahrungen in Ihre Festrede mit ein. [60 BE]	[TH, 2014, eA (1)]
T2	Entwerfen Sie auf der Grundlage der Materialien einen Essay, in dem Sie sich mit der Botschaft Jesu auseinandersetzen.	[TH, 2014, eA (2)]
T3	Zeigen Sie auf, welche Rolle christliche Maßstäbe für Ihre Entscheidungen spielen.	[TH, 2014, eA (2)]
T4	Prüfen Sie die Position der evangelischen Kirche zur Präimplantationsdiagnostik (Material 1). Entwerfen Sie einen Blog in Form eines Kommentars als Ihre persönliche Reaktion auf diese Veröffentlichung. Nutzen Sie zur Erstellung Ihres Internetbeitrages das Material 1. Beziehen Sie das Material 2 sowie geeignete Texte aus der Bibel in Ihre Betrachtungen ein. ⁴ [M2 = Text über »Designerbaby« → Genmanipulation]	[TH, 2015, eA (2)]
T5	Entwerfen Sie eine Trauerrede oder einen Essay zum Thema Sterben und Tod. Setzen Sie sich dabei mit der Theodizeefrage auseinander und zeigen Sie Perspektiven im Umgang mit Leid auf.	[TH, 2016, eA (1)]
T6	Erläutern Sie unter Zuhilfenahme mindestens zweier selbstgewählter Beispiele, inwiefern das Freiheitsverständnis Luthers heute noch Relevanz besitzt.	[TH, 2017, eA (1)]

⁴ Trotz der in der Formulierung enthaltenen zwei Operatoren wird diese Aufgabe als eine gezählt, da sich die Erstellung des Blogs mit der geforderten persönlichen Positionierung unmittelbar auf die Position der evangelischen Kirche bezieht, die zuvor geprüft werden soll.

Fortsetzung

Bez.	Aufgabenstellung	Signatur
T7	[3] Interpretieren Sie – auch mit Hilfe Ihrer Kenntnisse vom christlichen Menschenbild und der evangelischen Sozialethik – das Gedicht »Stufen« von Hermann Hesse (Material 2) unter der Fragestellung, wie sich Freiheit im Leben verantwortlich gestalten lässt.	[TH, 2017, eA (1)]
T8	Lässt sich mit der Bergpredigt Politik gestalten? Entwerfen Sie eine Antwort auf diese Frage, indem Sie die Textauszüge aus der Bergpredigt (Mt 5, 1–16 und Mt 5, 38–48) analysieren und die Problematik anhand von frei gewählten Beispielen aus Politik und Gesellschaft erörtern.	[TH, 2017, eA (2)]
T9	Vergleichen Sie die vorliegenden persönlichen Bekenntnisformulierungen (Material 1–4) mit dem apostolischen Glaubensbekenntnis. Zeigen Sie drei theologische Differenzen auf und nehmen Sie Stellung zu deren Bedeutung für den christlichen Glauben.	[TH, 2018, eA (1)]
T10	Beurteilen Sie den Wert von Gottesbeweisen für den Gottesglauben.	[TH, 2018, eA (2)]
T11	Entwerfen Sie einen offenen Brief, in dem Sie das in dem »Glaubensbekenntnis« des Textbeispiels enthaltene Gottesbild darstellen. dieses mit dem Ihnen bekannten biblischen Gottesbild vergleichen. sich persönlich mit den Gottesbildern auseinandersetzen.	[TH, 2019, eA (1)]
N1	3. Nehmen Sie in Form eines Beitrags für die Internetseite »evangelisch.de« Stellung zu Hintzes Position.	[NI, 2014, gA (1)]
N2	3. Nehmen Sie dazu Stellung, ob und inwiefern der Ansatz von Anja Middelbeck-Varwick für einen Dialog zwischen Christen und Muslimen tragfähig ist.	[NI, 2014, gA (2)]
N3	4. Setzen Sie sich unter Berücksichtigung ausgewählter Aspekte des Textes von Klaus-Peter Jörns mit dem Titelblatt des »SPIEGEL« (M2) auseinander. [Titelbild mit der Frage »Wer hat den stärkeren Gott?« – Zwei Hände, die mit den Zeigefingern ihre Kräfte messen]	[NI, 2014, eA (1)]
N4	3. Beurteilen Sie, inwieweit der vorliegende Textauszug eine altersgemäße und theologisch angemessene Darstellung der Theodizee-Frage ist.	[NI, 2015, gA (2)]

Fortsetzung

Bez.	Aufgabenstellung	Signatur
N5	4. Setzen Sie sich unter Bezugnahme auf ein oder mehrere konkrete Beispiele von Leiderfahrungen mit der Position Gre-shakes auseinander.	[NI, 2015, eA (1)]
N6	4. Nehmen Sie in einem Brief an den Vorstand der ACK Bayern einen begründeten Standpunkt zu deren Aufruf ein (M2). Be-ziehen Sie in Ihre Ausführungen ausgewählte Aspekte aus dem Text von Petra Bahr ein.	[NI, 2015, eA (2)]
N7	3. Entwerfen Sie einen Brief an Nikolaus Schneider zu dessen Sicht des Menschen.	[NI, 2016, gA (1)]
N8	3. »Die Religionskritik weist zu Recht auf das in vieler Hinsicht katastrophale Versagen der Religionen hin. Aber Religionskritik kann Religion nicht ersetzen [...]« (Z. 19–20) Nehmen Sie zu dieser Aussage vor dem Hintergrund eines reli-gionskritischen Konzeptes Stellung.	[NI, 2016, gA (2)]
N9	4. Setzen Sie sich mit Hubers Freiheitsverständnis auseinander.	[NI, 2016, eA (1)]
N10	3. Setzen Sie sich mit der Position Steffenskys auseinander.	[NI, 2017, gA (1)]
N11		[NI, 2017, eA (2)]
N12	3. Prüfen Sie, inwiefern das Konzept der »prophetischen Kirche« (Z. 35 ff.) Perspektiven für eine zukunftsfähige Kirche bietet.	[NI, 2017, gA (2)]
N13	3. Setzen Sie sich mit der Position des Leserbriefes (M 2) aus-einander.	[NI, 2018, gA (1)]
N14	3. Beurteilen Sie, ob Missalla überzeugende Antworten auf reli-gionskritische Einwände gibt.	[NI, 2018, gA (2)]
N15	4. Setzen Sie sich mit der Position Hubers auseinander.	[NI, 2018, eA (1)]
N16	4. Setzen Sie sich in einem Brief an Abdel-Samad damit aus-einander, in welcher Form uns insbesondere in der Folge der Aufklärung die Religionen in die Zukunft begleiten können.	[NI, 2018, eA (2)]
N17	3. Entwerfen Sie einen Kommentar zu Käßmanns friedens-ethischer Position.	[NI, 2019, gA (1)]
N18	3. Überprüfen Sie, ob und inwiefern Schneiders Aussagen zu Gottesbildern eine überzeugende Antwort auf klassische reli-gionskritische Einwände enthalten.	[NI, 2019, gA (2)]
N19	4. Entwerfen Sie einen Kommentar zu Käßmanns friedens-ethischer Position.	[NI, 2019, eA (1)]
N20	4. Setzen Sie sich mit Mettes Aussagen über die Versuchung (Z. 28 bis Ende) auseinander.	[NI, 2019, eA (2)]

Anhang IV: Urteilsbildungen mit Perspektivübernahme

Aufgabenstellungen, welche eine Urteilsbildung aus einer anderen Perspektive fordern, wurden aus dem Untersuchungskorpus ausgeschlossen (vgl. Kap. 5.2.3).

Aufgabenstellung	Signatur
4. In einer Online-Debatte darüber, was nach dem Tod mit Profilen und Konten bei sozialen Netzwerken im Internet geschehen soll, findet sich folgender Beitrag: »So lange dein Profil aktiv ist, wirst du niemals sterben.« [...] 4.2 Beurteilen Sie diese Vorstellung aus christlicher Perspektive! Gehen Sie dabei auch auf neutestamentliche Aussagen ein! [20 BE]	[BY, 2015 (1)]
4. »Zum Wesen des Protestantismus [gehört] die Vielfalt und es [gibt] nur auf wenige politische Fragen ethisch eindeutige Antworten.« (Z. 57–59) Erörtern Sie eine ethische Fragestellung Ihrer Wahl aus christlicher Perspektive! Gehen sie dabei auch auf die von Graf angesprochene Vielfalt und Uneindeutigkeit ein! [25 BE]	[BY, 2015 (4)]
4 Erörtern Sie eine medizinethische Fragestellung aus christlicher Perspektive! Beziehen Sie dabei Gedanken aus Material 1 und 2 ein!	[BY, 2016 (1)]
4 »Es gingen Bilder um die Welt, die sich nicht mehr aus dem Gedächtnis auslöschen lassen.« (Material 1, Zeile 2 f.) Ein Foto, das im September 2015 viele Menschen beschäftigte, zeigt ein totes Flüchtlingskind. Über die Veröffentlichung des Bildes wurde in den Medien kontrovers diskutiert (Materialien 2–4). Entwerfen Sie einen Leserbrief an eine der in den Materialien 2 und 4 genannten Zeitungen! Nehmen Sie darin zum Umgang dieser Zeitung mit dem Foto aus der Sicht einer christlichen Medienethik Stellung! Beziehen Sie die Materialien 1–4 mit ein! [30 BE]	[BY, 2016 (3)]
Ihre Schule will am Projekt »Netzgänger 3.0« (Material 2) teilnehmen. Bei einem Informationsabend für Eltern der Unterstufe soll die Notwendigkeit des Projekts aus unterschiedlichen Perspektiven beleuchtet werden. Entwerfen Sie für diesen Anlass als zukünftiger Multiplikator (Peer) einen Kurzvortrag, in dem Sie anhand einer frei gewählten Problemstellung der Medienethik die Durchführung des Projekts aus evangelischer Sicht begründen! Berücksichtigen Sie dabei zentrale Gedanken des Textes (Material 1)! [30]	[BY, 2019 (1)]

Fortsetzung

Aufgabenstellung	Signatur
4 »Erschaffe dich neu!« (Material 1, Z. 33). Stellen Sie ein selbstgewähltes Beispiel für das moderne Streben nach Selbstoptimierung dar und erörtern Sie davon ausgehend Möglichkeiten und Grenzen dieses Strebens aus christlicher Sicht! [25]	[BY, 2019 (2)]
5 Ihre Schule plant unter dem Titel »Der Mensch: In Verantwortung für die Welt!?« einen Aktionstag zum Thema Nachhaltigkeit. Dazu wird eine Seite auf der Schulhomepage gestaltet, auf der aus Sicht der einzelnen Fächer dargestellt wird, warum diesen das Thema ein Anliegen ist. Entwerfen Sie für das Fach Evangelische Religionslehre den entsprechenden Textbeitrag! Beziehen Sie dabei Gedanken aus dem Text (9 f.) mit ein! [30 BE]	[BY, 2019 (3)]
4 In einem religiös-weltanschaulich neutralen Staat »leben Menschen gemäß ihren durchaus unterschiedlichen religiösen oder sonstigen Überzeugungen« (Z. 8–10). Begründen Sie aus christlicher Sicht die Spannung zwischen Ausschließlichkeit und Beliebigkeit, die sich hinsichtlich des eigenen Wahrheitsanspruchs in der Auseinandersetzung mit anderen Religionen und Weltanschauungen ergibt! [15]	[BY, 2019 (4)]
[3.] Entwerfen Sie eine Antwort auf den Forumsbeitrag (Material 1) aus christlicher Perspektive. Nehmen Sie darin Stellung zu den aufgeworfenen Fragen unter Einbeziehung der Materialien 2–5, beziehen Sie dabei ihr Wissen mit ein und begründen Sie ihre [sic!] Argumentation.	[TH, 2015, eA (1)]

Anhang V: Gedicht »Stufen« von Hermann Hesse

Stufen

Wie jede Blüte welkt und jede Jugend
Dem Alter weicht, blüht jede Lebensstufe,
Blüht jede Weisheit auch und jede Tugend
Zu ihrer Zeit und darf nicht ewig dauern.
Es muß das Herz bei jedem Lebensrufe
Bereit zum Abschied sein und Neubeginne,
Um sich in Tapferkeit und ohne Trauern
In andre, neue Bindungen zu geben.
Und jedem Anfang wohnt ein Zauber inne,
Der uns beschützt und der uns hilft, zu leben.

Wir sollen heiter Raum um Raum durchschreiten,
An keinem wie an einer Heimat hängen,
Der Weltgeist will nicht fesseln uns und engen,
Er will uns Stuf um Stufe heben, weiten.
Kaum sind wir heimisch einem Lebenskreise
Und traulich eingewohnt, so droht Erschlaffen;
Nur wer bereit zu Aufbruch ist und Reise,
Mag lähmender Gewöhnung sich entrafen.
Es wird vielleicht auch noch die Todesstunde
Uns neuen Räumen jung entgegensenden,
Des Lebens Ruf an uns wird niemals enden ...
Wohlan denn Herz, nimm Abschied und gesunde!

HERMANN HESSE, Stufen, in: DERS., Gesammelte Werke. Erster Band. Suhrkamp Verlag, Frankfurt am Main 1987, 119. Das Gedicht schrieb Hesse nach langer Krankheit im Jahre 1941, ursprünglich mit dem Titel »Transzendieren«.



David Käbisch |
Johannes Wischmeyer (Hrsg.)

Wind of Change?

**»1968« und »1989« in der ost- und
westdeutschen Religionspädagogik**

Studien zur Religiösen Bildung (StRB) | 21

224 Seiten | Hardcover | 15,5 x 23 cm
ISBN 978-3-374-06442-7
EUR 48,00 [D]

»1968« und »1989« – wie in einem Brennglas bündeln diese beiden Daten den gesellschaftlichen Wandel. Dies gilt nicht zuletzt für die Rahmenbedingungen und Inhalte religiöser Erziehung und Sozialisation. Religionspädagogik und kirchliche Bildungspolitik haben sich im Gefolge von »'68« und »'89« massiv verändert. Emanzipation, Politisierung und ein Mehr an gesellschaftlicher Verantwortung prägten Diskurse und Praxen religiöser Bildung. Die Beiträge des Tagungsbands zeigen dabei eine bemerkenswerte Ähnlichkeit der Leitideen und Ziele in den jeweiligen Umbruchkontexten. Durch die übergreifende Betrachtung des Umbruchgeschehens zwischen 1968 und 1989 wirft der Band neue Schlaglichter auf die Verflechtungen ost- und westdeutscher sowie evangelischer und katholischer Religionspädagogik in der Reformära.



**EVANGELISCHE VERLAGSANSTALT
Leipzig** www.eva-leipzig.de

Tel +49 (0) 341/ 7 11 41 -44

shop@eva-leipzig.de

Mit Urteilskompetenz ist das Ziel verbunden, Lernende zu einer mündigen Auseinandersetzung mit ihrer Lebenswelt und zur gesellschaftlichen Partizipation zu befähigen. Doch wann ist jemand »urteilsfähig« und was wird geprüft und bewertet, wenn in Prüfungssituationen Urteilsfähigkeit gefordert wird?

Die Studie analysiert schriftliche Abituraufgabenstellungen und die dazugehörigen Erwartungshorizonte aus den Bundesländern Bayern, Niedersachsen und Thüringen. Sie greift damit einen hochaktuellen Diskurs über Vergleichbarkeit und Transparenz in der Bewertung komplexer Leistungen auf, der von fachübergreifendem Interesse ist. Resümierend werden konkrete Vorschläge formuliert, die sowohl für die Weiterentwicklung bildungsadministrativer Vorgaben als auch für die unterrichtliche Praxis zentral sind.

