

Sabine Krause/Roumiana Nikolova/Henning Schluß/Thomas Weiß/Joachim Willems

Kompetenzerwerb im evangelischen Religionsunterricht

Ergebnisse der Konstruktvalidierungsstudie der DFG-Projekte RU-Bi-Qua/KERK

Zusammenfassung: Wie in allen Fachdidaktiken gibt es auch im Bereich des Religionsunterrichts Bestrebungen, die fachspezifische Kompetenz so zu beschreiben, dass sie empirischen Erhebungen zugänglich wird. Der vorliegende Text stellt das Konzept der DFG-geförderten Projekte „Qualitätssicherung und Bildungsstandards für den Religionsunterricht an öffentlichen Schulen, am Beispiel des Evangelischen Religionsunterrichts“ und „Konstruktion und Erhebung von religiösen Kompetenzniveaus im Religionsunterricht am Beispiel des evangelischen Religionsunterrichts“ (RU-Bi-Qua/KERK) vor, die in der Kooperation von Erziehungswissenschaftlern, Religionspädagogen und einer Empirikerin ein Konzept religiöser Kompetenz und Aufgaben zu ihrer empirischen Testung entwickelt haben. Die Ergebnisse der Konstruktvalidierungsstudie werden hier erstmals diskutiert. Sie bestätigen den gewählten Ansatz in zentralen Punkten, zeigen zugleich jedoch auch die Richtung der notwendigen Weiterarbeit an.

1. Einleitung

„Der Religionsunterricht ist in den öffentlichen Schulen mit Ausnahme der bekenntnisfreien Schulen ordentliches Lehrfach. Unbeschadet des staatlichen Aufsichtsrechtes wird der Religionsunterricht in Übereinstimmung mit den Grundsätzen der Religionsgemeinschaften erteilt.“ Diese beiden Sätze aus dem Grundgesetz (Art. 7 Abs. 3 GG) umreißen die Stellung des Religionsunterrichts im deutschen Schulwesen.

Für die Bundesländer Berlin und Brandenburg gilt dieser Artikel des Grundgesetzes allerdings nicht (Berlin) oder seine Geltung ist umstritten (Brandenburg).¹

In allen Bundesländern jedoch wird das schulische Fach Religion „in Übereinstimmung mit den Grundsätzen der Religionsgemeinschaften erteilt“. Da es keine Religion „an sich“ gibt, sondern nur konkrete Religionen, und da es dem Staat nicht zukommt, zu bestimmen, was deren Grundsätze sind, ist er überall auf die Kooperation mit den anerkannten Religionsgemeinschaften angewiesen.² Die Benennung der Bezugskonfession

1 Während die sog. „Bremer Klausel“ (Art. 141 GG) für Berlin Art. 7 Abs. 3 außer Kraft setzt und der Religionsunterricht hier ein freiwilliges und im Zeugnis nicht benotetes Fach ist, das zwar in den Räumen der Schule unterrichtet wird, dessen Gestaltung aber allein in der Verantwortung der Kirchen bzw. Religions- oder Weltanschauungsgemeinschaften liegt, berief sich das Land Brandenburg im Rechtsstreit um LER ohne letztgültigen Entscheid auf diese Klausel. Dennoch ist die Situation des RU in Brandenburg eine besondere, da er am ehesten mit dem Status eines Ersatzfaches für LER zu vergleichen ist, das als Regelfach gilt.

2 Zum Status des Religionsunterrichts in den einzelnen Bundesländern am aktuellsten vgl. den Schwerpunkt „Regionalisierung von Religionsunterricht und Religionspädagogik“ in: www.theo-web.de 2006, H. 2 und 2007, H. 1.

(meist evangelisch oder katholisch) macht für alle Beteiligten transparent, von welcher Position aus das Fach, das nicht ‚neutral‘ unterrichtet werden kann, gelehrt wird. Trotz dieser institutionellen Besonderheiten ist der Religionsunterricht insofern ein Unterrichtsfach wie andere Fächer auch, als es in ihm um Wissens- und Kompetenzerwerb geht und die Leistungen (nicht Gesinnungen, Einstellungen oder Bekenntnisse) bewertet werden. Da es sich um ein schulisches Fach handelt, muss den Schülerinnen und Schülern die Möglichkeit eröffnet werden, eigene Stellungnahmen zu den Unterrichtsgegenständen zu finden und zu reflektieren.³ In Zeiten einer Umstellung des Schulwesens von der Input- zur Outputsteuerung (Klieme et al. 2003, S. 11f.) steht deshalb auch der evangelische Religionsunterricht vor der Aufgabe, seinen Beitrag zur allgemeinen Bildung so auszuweisen, dass deutlich wird, welche fachspezifischen Kompetenzen in diesem Unterricht erworben werden.

Im Rahmen zweier von der DFG geförderter Projekte – RU-Bi-Qua⁴ und KERK – erhebt und beschreibt seit Februar 2006 eine an der Humboldt-Universität zu Berlin gegründete interdisziplinäre Gruppe von Allgemeinen Erziehungswissenschaftlern, Religionspädagogen und einer erziehungswissenschaftlichen Empirikerin religiöse Kompetenz als fachspezifische Kompetenz. Der Begriff „Religion“ wird im Projekt als diskursiver Tatbestand definiert (Matthes 1992). Ein solches Verfahren bietet die Chance, dass die entwickelten fachspezifischen Kompetenzen nicht nur für einen unter vielen möglichen Religionsbegriffen gelten, sondern dass auch die Beschreibung dessen, was als Religion zu verstehen ist, nicht als eine Voraussetzung des Unterrichts, sondern als einer seiner Gegenstände verstanden werden kann. Das auf religionspädagogischem und bildungstheoretischem Hintergrund entwickelte Konzept religiöser Kompetenz wird in den Bundesländern Berlin und Brandenburg empirisch getestet und auf der so gewonnenen Datengrundlage werden Dimensionen und Niveaus religiöser Kompetenz unterschieden.

Die besondere Situation des Religionsunterrichts in Berlin und Brandenburg hat einerseits eine relativ hohe Fluktuation und eine im Vergleich zu den alten Bundesländern im Durchschnitt geringere Teilnahme am Religionsunterricht zur Folge. Andererseits kommt der Freiwilligkeit des Besuchs durch diese Konstruktion (mit Ausnahme

3 In diesem Sinne betont die Evangelische Kirche in Deutschland, dass der konfessionelle Religionsunterricht „im Lichte von Artikel 4 GG, des Rechts auf Religionsfreiheit, auszulegen ist. Er hat der ‚Sicherung der Grundrechtsausübung durch den einzelnen‘ zu dienen, dem einzelnen Kind und Jugendlichen. Sie sollen sich frei und selbständig religiös orientieren können.“ (EKD 1994, S. 11)

4 Der Titel des DFG-Projekts lautet: „Qualitätssicherung und Bildungsstandards für den Religionsunterricht an öffentlichen Schulen, am Beispiel des Evangelischen Religionsunterrichts (RU-Bi-Qua).“ Dieser Titel ist insofern irritierend, als das Projekt nicht auf die Festlegung von Bildungsstandards ausgerichtet ist, sondern religiöse Kompetenzen erheben möchte. Dies wurde im reformulierten Titel des gegenwärtig laufenden, ebenfalls von der DFG geförderten, Projekts berücksichtigt: „Konstruktion und Erhebung von religiösen Kompetenzniveaus im Religionsunterricht am Beispiel des evangelischen Religionsunterrichts. (KERK)“ (Benner 2007; Benner/Krause/Nikolova et al. 2007; Schieder 2007; Nikolova/Schluß/Weiß/Willems 2007).

der kirchlichen Schulen, die auch in den Test einbezogen sind) eine höhere Bedeutung zu. Die Erforschung der Effekte, die diese organisatorische Besonderheit im Vergleich zum Religionsunterricht als ordentlichem Lehrfach nach GG Art. 3 in Bezug auf die Ausprägung religiöser Kompetenz hat, sind ebenfalls Gegenstand der Untersuchungen des Projekts, das sich um eine Vergleichserhebung in anderen Bundesländern bemüht.

2. Das Konzept fachspezifischer religiöser Kompetenz

Das hier vertretene Konzept religiöser Kompetenz beschränkt sich auf den testbaren Teil unterrichtlicher Bemühungen, also in erster Linie auf die kognitiven Leistungen der Schülerinnen und Schüler im religiösen Bereich.⁵ Unter religiöser Kompetenz wird dabei nicht nur die Erfassung von Wissensbeständen⁶ verstanden, sondern die Fähigkeit, diese Wissensbestände und andere Inhalte des Religionsunterrichts kontextbezogen zu interpretieren, verschiedene Weltbezüge (wie Religion, Ethik, Politik, Ökonomie) und deren eigene Logiken zu unterscheiden und zu dem jeweiligen Gegenstand in Beziehung setzen zu können. Damit deckt das Konzept nicht alles ab, was theologisch oder in einem spirituellen Sinn als religiöse Kompetenz beschreibbar wäre. In dieser Beschränkung der fachspezifischen religiösen Kompetenz auf diesen eng abgesteckten Bereich im weiten Feld der Religion(en) ist das Konzept anschlussfähig an andere fachbezogene Konzeptualisierungen und Testungen fach- und domänenspezifischer Kompetenzen. Ziel der erhobenen religiösen Kompetenz ist über das Wissen hinaus ein Können, d.h. Schülerinnen und Schüler sind dann religiös kompetent, wenn sie nicht nur Wissen wiedergeben können, sondern mit diesem Wissen darüber hinaus reflexiv umgehen können.

Religiöse Kompetenz als fachspezifische Kompetenz ist im vorliegenden Modell bezogen auf drei Gegenstandsbereiche:

- 1) Bezugsreligion/-konfession des Unterrichts: Anstelle des vielfach gebräuchlichen Begriffs der „eigenen Religion“ wird der Begriff „Bezugsreligion“ verwendet, weil keineswegs mehr vorausgesetzt werden kann, dass die Bezugsreligion/-konfession des Unterrichtsfaches immer die „eigene Religion“ der Schülerinnen und Schüler ist.
- 2) Andere Religionen: Religiöse Kompetenz muss sich auch auf andere Konfessionen und Religionen erstrecken. Ihre interkonfessionelle und interreligiöse Perspektive geht zwar von der Bezugsreligion aus, verlangt jedoch, dass die Schülerinnen und Schüler lernen, zwischen verschiedenen Religionen und Konfessionen den Wechsel von Perspektiven einzutüben und zu reflektieren. Um diesen Kompetenzbereich zu

5 Der mit diesem Modell überprüfte Kompetenzbegriff ist damit enger als der von Klieme und Weinert, welcher sich z.B. auch auf die Motivation erstreckt, die von RU-Bi-Qua/KERK nicht erhoben wird (Klieme 2003, S. 59).

6 Das Projekt RU-Bi-Qua/KERK erhebt auch grundlegende Wissensbestände, die als religionskundliche Kenntnisse (vgl. Punkt 3) ausgewiesen sind.

fördern, muss der Religionsunterricht auch Quellen, Zeugnisse und Rituale aus anderen Konfessionen und Religionen thematisieren und interreligiösen Austausch ermöglichen.

- 3) Religiöse Aspekte in außerreligiösen Bereichen: Religiöse Kompetenz bezieht sich auch auf Bereiche der Gesellschaft und Kultur, die zentral nach Maßgabe anderer Handlungslogiken als religiösen (z.B. ökonomischen, ethischen oder politischen) interpretiert werden können. Ihre Bezüge zum Religiösen hängen damit zusammen, dass sich einzelne Phänomene selten exklusiv einem Bereich zuordnen lassen, sondern auch unter anderen Blickwinkeln thematisiert werden können. So werden sportliche Großereignisse nicht nur unter der Logik des sportlichen Wettbewerbs, sondern auch unter ökonomischen Leitfragen oder religiösen Kriterien sinnvoll verstanden werden können. Aufgabe des Religionsunterrichts ist es hier, die eventuell vorhandenen religiösen Ursprünge oder Konnotationen zu thematisieren und in Bezug auf die „anderen“ Bereiche im öffentlichen Raum zu reflektieren.

Religiöse Kompetenz als fachspezifische Kompetenz teilt sich in eine *religiöse Deutungskompetenz* (a) und eine *religiöse Partizipationskompetenz* (b).

Zu a) Phänomene, die religiöse Bezüge aufweisen, als solche wahrzunehmen und zu erfahren sowie diese Wahrnehmungen und Erfahrungen durch im Unterricht erworbene Kenntnisse zu erweitern und in ihren vielfältigen Bedeutungshorizonten reflexiv beurteilen und verarbeiten zu können, wird als *religiöse Deutungskompetenz* verstanden.

Zu b) *Religiöse Partizipationskompetenz* wird als ein Können verstanden, das es Schülerinnen und Schülern erlaubt, an einer Handlung teilzuhaben und einen Handlungsverlauf aus verschiedenen Perspektiven nachzuvollziehen zu können. Religiöse Partizipationskompetenz ermöglicht es Schülerinnen und Schülern auf diese Weise, zu religiösen Sachverhalten oder Phänomenen mit religiösen Aspekten individuell, gemeinsam und öffentlich Stellung zu nehmen. Diese Betonung der Stellungnahme ist deshalb bedeutsam, weil durch diese Definition religiöser Partizipationskompetenz dem Missverständnis gewehrt werden soll, dass es Aufgabe des Religionsunterrichts sei, Heranwachsende zu missionieren und für ein bestimmtes Bekenntnis zu gewinnen. Das hier vorgestellte Verständnis zur religiösen Partizipationskompetenz orientiert sich am Konzept der Problemlösungskompetenz (Klieme/Funke/Leutner et al. 2001, S. 179ff.).

| Gegenstandsbereiche | Religiöse Deutungskompetenz | Religiöse Partizipationskompetenz |
|-------------------------|-----------------------------|-----------------------------------|
| Bezugsreligion | | |
| Andere Religionen | | |
| Außerreligiöse Bereiche | | |

Die religiöse Kompetenz – differenziert in religiöse Deutungs- und religiöse Partizipationskompetenz – erstreckt sich dabei über drei Ebenen, die zueinander in einem korrespondierenden und sich wechselseitig erweiternden Verhältnis stehen. Auf der ersten Ebene sind Erfahrungen mit religiösen Phänomenen oder Partizipationserfahrungen in religiösen Kontexten zu nennen. Erfahrungen sind Voraussetzung eines Unterrichts, der die Aufgabe hat, Erfahrungen reflexiv, wissenschaftspropädeutisch und wissenschaftlich zu erweitern. Erfahrungen selbst sind allerdings nicht bewertbar. Deshalb werden Erfahrungen der Schüler in einem Fragebogen als einem besonderen Teil des Testinstruments erfasst. Die Erarbeitung eines Fragebogens zur Lehrerperspektive steht noch aus. Die zweite Ebene bezieht sich auf die Kenntnisse im Bereich des Religiösen, die fachspezifisch zu vermitteln sind. Diese Kenntnisse werden gesondert erhoben und können somit zur dritten Ebene, den hermeneutischen Fähigkeiten auf Seiten der religiösen Deutungskompetenz und der Reflexion und Stellungnahme zu religiösen Partizipationserfahrungen und -möglichkeiten in Beziehung gesetzt werden.⁷

| | |
|---------------------------------|---|
| Religiöse Deutungskompetenz | Religiöse Partizipationskompetenz |
| Erfahrungen mit rel. Phänomenen | Partizipationserfahrungen in religiösen oder religiös interpretierbaren Feldern |
| Religionskundliche Kenntnisse | |
| Hermeneutische Fähigkeiten | Reflexion und Stellungnahme zu religiösen Partizipations-Erfahrungen und -möglichkeiten |

Unter Beachtung religionspädagogischer, bildungstheoretischer und empirischer Anforderungen wurden im Projekt Testaufgaben zur Erhebung religiöser Kompetenz in all den dargestellten Bezügen entwickelt.⁸ In mehreren Pretestungen konnten die Aufgabenstellungen und die Antwortvorgaben verfeinert bzw. angepasst werden. Erst nach den erfolgten Testungen⁹ war es möglich, für die religiöse Deutungskompetenz Schwierigkeitsniveaus zu formulieren und eine erste Interpretation von Beispielitems vorzulegen (Benner/Krause/Nikolova et al., 2007, bes. S. 147ff.).

3. Ergebnisse der Konstruktvalidierungsstudie

Im Folgenden werden Ergebnisse aus der Pilotuntersuchung vorgestellt, die im Rahmen des Aufgabenentwicklungsprozesses von RU-Bi-Qua/KERK an Berliner und Branden-

⁷ Das Konzept wurde im Laufe des Forschungsprojektes in der Auseinandersetzung mit den empirischen Befunden weiterentwickelt (Nikolova/Schluss/Weiß/Willems 2007).

⁸ Die Entwicklung von Multiple-Choice-Aufgaben zur empirischen Erhebung der Deutungskompetenz hat sich dabei bislang als prinzipiell geeignet erwiesen.

⁹ Im Zuge der Aufgabenentwicklung gab es mehrere kleinere Stichproben mit Teilnehmerzahlen zwischen 30 und 102 Schüler/innen im Zeitraum von Juni 2006 bis Februar 2007.

burger Schulen im Juni 2007 durchgeführt wurde. Diese Pilotstudie sollte Informationen zur Zuverlässigkeit, Validität und Einsetzbarkeit der entwickelten Testaufgaben erbringen sowie der empirischen Validierung des generierten Kompetenzmodells an einer größeren Stichprobe dienen. Sie bezieht sich mit der Deutungskompetenz auf die linke Spalte der Tabelle 2 und hielt Testaufgaben für die Bereiche „religionskundliche Kenntnisse“ und „hermeneutische Fähigkeiten“ bereit.

Die Darstellung und Diskussion der Befunde orientiert sich im Folgenden an fünf Aspekten:

- 1) Dimensionalitätsprüfung und interne Konsistenz der getesteten Items (Item-Fit-Analysen).
- 2) Aufgabenschwierigkeiten und Fairness der Items.
- 3) Konstruktvaliditätsüberprüfung bezogen auf Geschlecht- und Schulformunterschied.
- 4) Sensitivität für außerschulische Einflüsse: Zusammenhänge mit religiöser Erfahrungen in der Familie sowie mit dem sozialen Hintergrund der Schülerinnen und Schüler.
- 5) Sensitivität für Einflüsse der schulischen Bildung: Mehrebenenanalytische Zerlegung der Varianzanteile.

Zuvor wird im Folgenden das Instrument, das Design und die Stichprobe der Pilotierungsstudie vorgestellt.

3.1 Design und Testinstrument

Grundlage der am Ende des Schuljahres 2006/2007 in zehnten Klassen durchgeführten Pilotuntersuchung bildeten 60 Testaufgaben, die sich nach mehreren Prepilottierungserhebungen mit kleineren Stichproben teststatistisch als geeignet erwiesen haben. Bezogen auf das theoretische Konzept sind dabei 31 der getesteten Items der Skala „religionskundliche Grundkenntnisse“ und 29 der Skala „hermeneutische Fähigkeiten“ zuzuordnen.

Zu den Items, die der Skala „religionskundliche Grundkenntnisse“ zugeordnet wurden, gehören beispielsweise Fragen wie: „Welche sind die fünf Säulen des Islam“ oder „Wie viele Evangelien stehen im neuen Testament?“ Zur Skala „hermeneutische Fähigkeiten“ gehört z.B. die Frage (zum im Testheft abgedruckten Psalm 23): „Welche Erfahrungen mag der Dichter des Liedes gemacht haben?“. Hier mussten die Testpersonen unterscheiden können, inwieweit die Dichtung, bezogen auf Gott, auf Überwachung, Vertrauen, Schutz oder ein Festmahl abzielt. Die Testung ergab, dass dieses Item dem leichten Schwierigkeitsniveau zuzuordnen ist. Ein Item, das sich als sehr anspruchsvoll herausstellte, steht im Zusammenhang mit dem Gleichnis „Von den Arbeitern im Weinberg“. Gefragt wurde nach dem im Text vorausgesetzten Verhältnis zwischen Lohn und Arbeit. Die länger Arbeitenden werden in ihrer Erwartung, einen höheren Lohn zu er-

halten, enttäuscht. Das Finden der richtigen Antwort setzt ein differenziertes Verständnis zwischen ökonomischer und religiöser Logik voraus, um diese Zugangsweisen im Lichte des Textes als unterschiedene interpretieren zu können. Eines der weiteren Items fragt dann z.B., was dieses Gleichnis über das Reich Gottes aussagt. Hier bestand die Schwierigkeit darin, den theologisch geprägten „terminus technicus“ (Reich Gottes) als ein nicht aus (Erwerbs-) Arbeit ableitbaren Anspruch in den entsprechenden Antwortangeboten identifizieren zu können.¹⁰

Zusätzlich wurden auch Fragen mit Hintergrundinformationen zum religiösen und sozialen Umfeld der Testpersonen sowie zu Schülereinschätzungen des Religionsunterrichts in Form eines Schülerfragebogens eingesetzt. Da es aus zeitlichen und finanziellen Gründen nicht möglich war, einen weiteren Test (beispielsweise Kognitiver Fähigkeitstest (KFT) oder Leseverständnistest) als externe Validierungsquelle in die Pilotuntersuchung aufzunehmen, sollten die Informationen aus dem Fragebogen für die Konstruktvalidierungsanalysen herangezogen werden. Gleichzeitig wurden dabei auch im Rahmen der Pilotstudie nicht nur die Testaufgaben validiert, sondern auch die Hintergrundfragen und die Einschätzungsskalen im Fragebogen erprobt. Damit ergaben sich für die Pilotierung zwei Bestandteile: ein Test bestehend aus 60 Testaufgaben und ein Fragebogenteil mit 19 Fragekomplexen.

Da dieses Testinstrument an einer Stichprobe von mindestens 500 Schülerinnen und Schülern aller Schulformen in Berlin und Brandenburg getestet werden und dabei die Testzeit den Rahmen einer Unterrichtsstunde nicht wesentlich überschreiten sollte, wurde in Anlehnung an PISA und LAU ein Multi-Matrix-Testdesign realisiert, das schulformspezifisch angelegt ist. So wurden zwei Testhefte vorbereitet: ein Gymnasialheft und ein Haupt-, Gesamt-, und Realschulheft, die mit einem Kern von 21 gemeinsamen Aufgaben (11 Kenntnisaufgaben und 10 Interpretationsaufgaben) verankert wurden. Die Gymnasiastinnen und Gymnasiasten bekamen insgesamt 41 Testaufgaben, während die Schüler anderer Schulformen insgesamt 40 Items zu bearbeiten hatten. Der Schülerfragebogen war für alle Getesteten gleich. Dabei wurden 40 Minuten für die Bearbeitung des Testteils und 5 bis 7 Minuten für die Bearbeitung des Fragebogenteils zur Verfügung gestellt.

3.2 Stichprobe

Insgesamt konnten 23 Schulen für die Pilottestung gewonnen werden. Fünf Brandenburger Schulen mit insgesamt 178 Schülerinnen und Schülern der Gymnasial-, Real- und Gesamtschulform und 18 Berliner Schulen der Gymnasial-, Real-, Gesamt- und Hauptschulform mit insgesamt 341 Zehnklässlerinnen und Zehnklässlern haben sich an der Pilottestung von Ende Mai bis Anfang Juni 2007 beteiligt. Dadurch konnte die angestrebte Mindeststichprobenzahl von 500 Schülerinnen und Schülern sogar leicht über-

¹⁰ Beispielaufgaben aller Schwierigkeitsbereiche sind vorgestellt und die Gründe für die Schwierigkeit interpretiert in: Benner/Krause/Nikolova et al. 2007.

troffen werden. Die Schülerinnen und Schüler im Test kamen zu 69% aus Gymnasien, zu 23% aus Realschulen, zu 5% aus Hauptschulen und zu 3% aus Gesamtschulen. In Bezug auf das Geschlecht der Testpersonen konnte eine Gleichverteilung von nahezu 50% für jedes Geschlecht realisiert werden. Bis zu 94% der Testpersonen sprechen nur oder meistens Deutsch zu Hause. Lediglich 6% gaben einen anderen Sprachhintergrund an. Gemessen an der elterlichen Berufsausbildung gaben 1% der Befragten an, dass ihre Eltern keine berufliche Ausbildung haben; 12% gaben eine Lehre oder Berufsaufbauschule, 26% eine Berufsfach- oder Fachschule, 28% der Getesteten einen Fachhochschulabschluss der Eltern und 33% einen Hochschulabschluss als höchste elterliche Ausbildung an.

3.3 Ergebnisse

3.3.1 Zur Dimensionalität, Item-Fit und Zuverlässigkeit

Die eingesetzten Testitems wurden mit dem einparametrischen Rasch-Modell einer multidimensionalen Skalierung unterzogen. Dabei konnten 2 der 29 Interpretationsaufgaben wegen Auswertungsfehlern nicht berücksichtigt werden. Insgesamt wurden somit aufgrund theoretischer Überlegungen 27 Aufgaben der Dimension „hermeneutische Fähigkeiten“ und 31 Aufgaben der Dimension „religionskundliche Grundkenntnisse“ zugeordnet. Die multidimensionale Skalierung ergab eine latente (d.h. messfehlerbereinigte) Korrelation zwischen den beiden Skalen von $r=0.85$. Somit ist nachgewiesen, dass die Ergebnisse beider Skalen nicht identisch sind und deshalb die zweidimensionale Skalierung theoretisch sinnvoll und empirisch begründet ist, um die Bezüge zwischen hermeneutischen Fähigkeiten und Kenntnissen untersuchen zu können. In Tabelle 3 sind die Ergebnisse der Item-Fit-Analysen (weighted MNSQ) und der Analysen zur internen Konsistenz (WLE Reliabilität) der beiden Skalen zusammengefasst.

| Tab. 3: Dimensionalität, Item-Fit und Reliabilität der Items | | |
|--|--|---------------------------------------|
| Dimensionalität, Item-Fit und Reliabilität der eingesetzten Items | | |
| Skala | Item-Fit Indizien Ziel: zw. 0.80 und 1.20 | WLE Reliabilität Ziel: ≥ 0.65 |
| Religionskundliche Grundkenntnisse 31 Items, davon haben 29 Items einen Infit zw. 0.87 und 1.1 | weighted MNSQ $\geq 0.87 \leq 1.20$ | 0.72 |
| Hermeneutische Fähigkeiten 27 Items, davon haben 26 Items einen Infit zw. 0.87 und 1.1 | weighted MNSQ $\geq 0.87 \leq 1.16$ | 0.68 |

Die Kennwerte sind aus teststatistischer Sicht als befriedigend bis gut zu betrachten. Der niedrigere Wert bei der Skala „hermeneutische Fähigkeiten“ resultiert vermutlich aus der Tatsache, dass die Aufgaben – insbesondere im oberen Schwierigkeitsbereich – in unterschiedlicher Weise das Interpretieren von religiösen Inhalten aus komplexeren

Textzusammenhängen im Lichte der Kenntnis verschiedener Urteils- und Handlungslogiken (z.B. ökonomischer, religiöser, politischer, ethischer) mit ansprechen.

3.3.2 *Zur Aufgabenschwierigkeit und Itemfairness*

Die mittlere Aufgabenlösungshäufigkeit beträgt jeweils 0.56 für die Kenntnisskala und 0.60 für die Interpretationsskala. Die Aufgaben der beiden Skalen vermögen somit sehr gut zwischen den Testpersonen zu differenzieren. Für beide Skalen ist die Verteilung der Kompetenzen auf Mittelwert 100 und Standardabweichung 30 normiert. Die Verteilungen der Personenwerte sind annähernd normal verteilt und variieren zwischen 12 und 177 Skalenpunkten für die Interpretationsskala sowie zwischen 16 und 207 Skalenpunkten für die Kenntnisskala. Hinsichtlich der Aufgabenschwierigkeit hat sich bei einem Mittelwert von 84 Raschskalenpunkten eine breite Aufgabenverteilung von 30 bis 157 Skalenpunkten für die Interpretationsskala ergeben. Die Aufgaben dieser Skala lassen sich inhaltlich 8 Schwierigkeitsbereichen zuordnen, aus denen es theoretisch und empirisch möglich sein wird, Kompetenzstufen ähnlich wie in der PISA-Lesekompetenzskala abzuleiten. Die Aufgabenschwierigkeiten der Kenntnisskala liegen bei einem Mittelwert von 88 Raschskalenpunkten auf einem breiten Spektrum zwischen 24 und 162 Skalenpunkten. Die Analyse der Anforderungsstruktur der Kenntnisaufgaben lässt im Gegensatz zu der Interpretationsskala keine theoretisch begründbare Definition von Kompetenzstufen zu. Die zunehmende Schwierigkeit der Aufgaben scheint von der Vertrautheit des Lernstoffes abhängig zu sein.

Wenn die Antwortwahrscheinlichkeit eines Items nicht nur von der Personenfähigkeit und der Aufgabenschwierigkeit abhängt, sondern auch beispielsweise vom Geschlecht oder dem ethnischen Hintergrund der Probanden, so kann man davon ausgehen, dass das Testverfahren keinen fairen Vergleich zulässt. Deshalb muss bei der Aufgabenentwicklung nicht nur die Zuverlässigkeit und der Item-Fit, sondern auch die Existenz von unerwünschten differenziellen Effekten von Aufgaben (der so genannten Differential Item Functioning) überprüft werden. Die Überprüfung möglicher differenzieller Effekte wurde im Hinblick auf das Geschlecht durchgeführt. Für beide Skalen wurden keine solche Effekte festgestellt. Die korrelativen Zusammenhänge zwischen den Skalenwerten und dem Geschlecht sowie dem ethnischen Hintergrund erbrachten schwache bis keine signifikanten Ergebnisse, was für eine gesicherte Fairness eines Kompetenzvergleichs auf der Basis des im Rahmen von RU-Bi-Qua/KERK entwickelten Testinstruments spricht.

3.3.3 *Konstruktvaliditätsüberprüfung bezogen auf Geschlecht- und Schulformunterschied*

Um die Konstruktvalidität der Testverfahren zu untersuchen wurden die Daten im Hinblick auf Geschlechts- und Schulformeffekte analysiert. Aus sozialkognitiver Sicht kann

davon ausgegangen werden, dass Mädchen gegenüber Jungen in diesem Alter einen kognitiven Entwicklungsvorsprung zeigen. Darüber hinaus sollten auch Gymnasiasten auf einem höheren Niveau liegen, da sie im Durchschnitt auch in anderen großen Untersuchungen ein höheres kognitives Leistungsvermögen aufweisen. Abbildung 3¹¹ und Abbildung 4 zeigen, dass sowohl Mädchen als auch Gymnasiasten erwartungskonform signifikant¹² höhere Durchschnittswerte auf beiden Skalen erreichen. Die Ergebnisse bestätigen damit die Konstruktvalidität der beiden Skalen.

3.3.4 *Sensitivität für außerschulische Einflüsse – religiöse Erfahrungen und sozialer Hintergrund*

Wenn das Testinstrument gut zwischen Testpersonen mit und ohne einem religiösen Hintergrund bzw. zwischen Schülerinnen und Schülern aus bildungsnahen und bildungsfernen Elternhäusern differenziert, sollten auch klare Unterschiede in den Mittelwerten der Vergleichsgruppen bestehen. Für diese Analysen wurde im Fragebogen nach der Erfahrung mit Religion in Familie und Kirchgemeinde sowie nach dem Buchbestand in der Familie gefragt. Auf der Basis von vier Items (Bsp.: „Meine Eltern haben mir von Gott erzählt.“) zum Vorkommen und zur Häufigkeit religiöser Erfahrungen in der Familie wurde eine entsprechende Skala gebildet (Cronbach's Alpha=.83). Abbildung 3 fasst die kompetenzspezifischen Unterschiede im Hinblick auf das religiöse Erfahrungsumfeld und Abbildung 4 die soziokulturelle Lernumwelt, indiziert durch den familiären Buchbesitz, zusammen.

Es zeigen sich deutliche Differenzen¹³ sowohl bei der „hermeneutischen Fähigkeit“ als auch in den „religionskundlichen Grundkenntnissen“ zwischen den Befragten mit und ohne regelmäßige religiöse Erfahrungen. Dabei sind die Unterschiede im Bereich der Kenntnisse am stärksten ausgeprägt. Ein ähnliches Befundmuster wird auch im Hinblick auf die soziokulturelle Lernumwelt identifiziert. Die Schülerinnen und Schüler, die einen überdurchschnittlichen Buchbestand angegeben haben (mehr als 100 Bücher zu Hause), erreichen mit Abstand¹⁴ eine höhere religiöse Kompetenz im Sinne unserer Testfragen. Mit beiden Skalen lassen sich die differenziellen Kompetenzmuster sehr gut erfassen. Die Ergebnisse bestätigen somit die Sensitivität des Testinstruments für Einflüsse der außerschulischen Lernumwelt.

11 HF=Hermeneutische Fähigkeiten; RGK=Religionskundliche Grundkenntnisse.

12 Auf der Skala „hermeneutische Fähigkeiten“ sind die Mittelwertunterschiede zwischen den Geschlechtern auf dem 1-Prozent-Niveau signifikant; auf der Skala „religionskundliche Grundkenntnisse“ auf dem 5-Prozent-Niveau signifikant. Mittelwertunterschiede zwischen den Gymnasiasten und den anderen Schulformen sind auf beiden Skalen auf dem 1-Prozent-Niveau signifikant.

13 Mittelwertunterschiede auf beiden Skalen sind auf dem 1-Prozent-Niveau signifikant.

14 Die Mittelwertunterschiede der Schülerinnen und Schüler mit überdurchschnittlichem Buchbestand sind auf beiden Skalen auf dem 1-Prozent-Niveau signifikant.

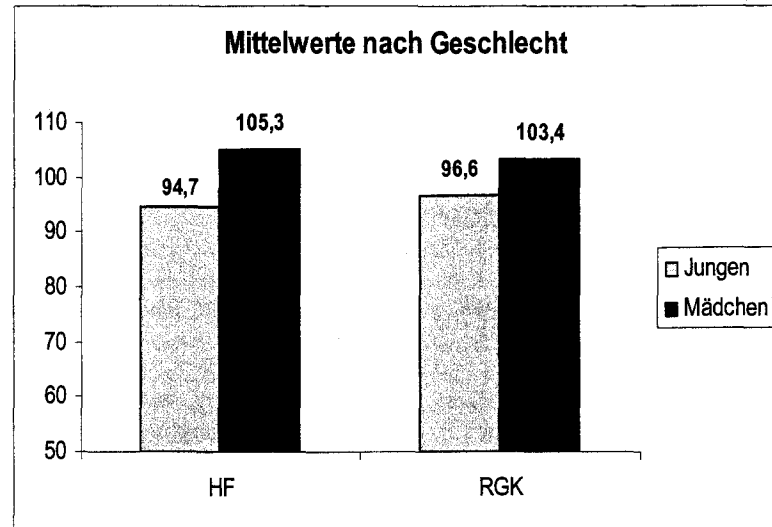


Abb. 1: Mittelwerte nach Geschlecht¹⁵

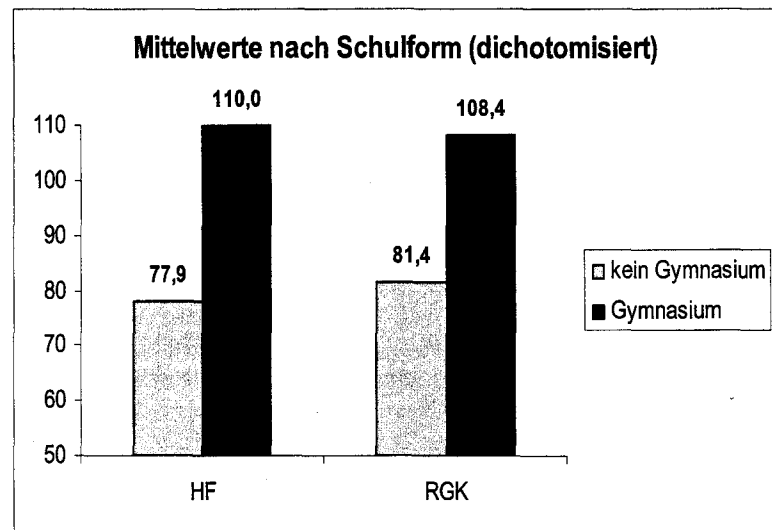


Abb. 2: Mittelwerte nach Schulform¹⁶

15 SD_{HF} Mädchen=29,1; SD_{HF} Jungen=30,0; SD_{RGK} Mädchen=28,6; SD_{RGK} Jungen=31,0.
 16 SD_{HF} Gymnasiasten=24,7; SD_{HF} Nicht Gymnasiasten=28,8; SD_{RGK} Gymnasiasten=27,7; SD_{RGK} Nicht Gymnasiasten=26,4.

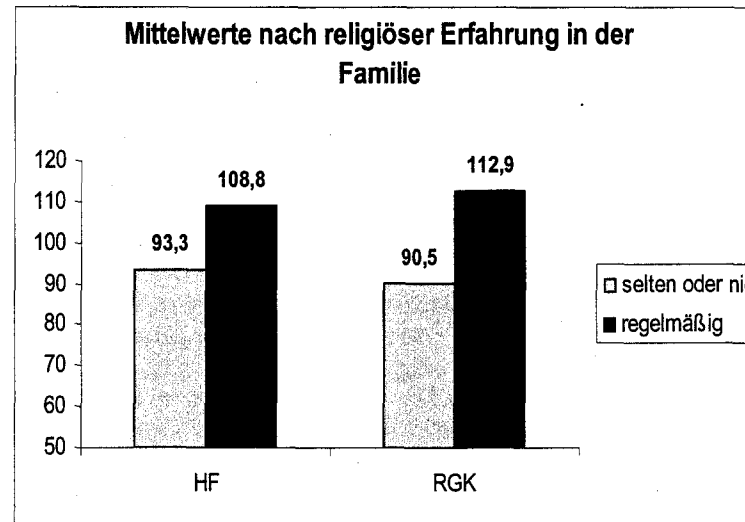


Abb. 3: Mittelwerte nach religiöser Erfahrung in der Familie¹⁷

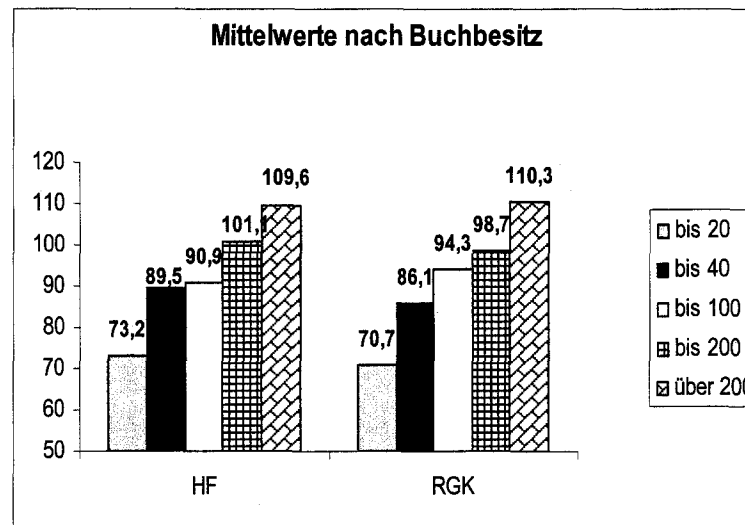


Abb. 4: Mittelwerte nach Buchbesitz in der Familie¹⁸

17 SD_{HF} Regelmäßige religiöse Erfahrung in der Familie=26,4; SD_{HF} Seltene religiöse Erfahrung in der Familie=31,1; SD_{RGK} Regelmäßige religiöse Erfahrung in der Familie=30,6; SD_{RGK} Seltene religiöse Erfahrung in der Familie=26,0.

18 SD_{HF} bis 20=27,1; SD_{HF} bis 40=30,0; SD_{HF} bis 100=27,5; SD_{HF} bis 200=25,6; SD_{HF} über 200=28,1; SD_{RGK} bis 20=23,7; SD_{RGK} bis 40=27,5; SD_{RGK} bis 100=24,9; SD_{RGK} bis 200=24,5; SD_{RGK} über 200=30,5.

3.3.5 Sensitivität für Einflüsse der schulischen Bildung – Mehrebenenanalytische Zerlegung der Varianzanteile

In welchem Ausmaß wird nun die religiöse Deutungskompetenz, hier als hermeneutische Fähigkeit und religionskundliche Grundkenntnisse erfasst, durch Effekte schulisch institutionalisierter Bildung beeinflusst? Inwieweit lassen sich schulische Effekte mit dem hier erprobten Testinstrument nachweisen? Zur Untersuchung dieser Fragen wurden die Daten einer mehrebenenanalytischen Modellierung unterzogen. In einem ersten Schritt wurde untersucht, welcher Teil der Kompetenzvarianz für beide Skalen überhaupt rechnerisch auf Schuleffekte zurückzuführen ist. In einem zweiten Schritt wurde nach der Kontrolle der individuellen Lernvoraussetzungen auch ein möglicher Schulförderungseffekt der Gymnasien analysiert.

Die Rechnungen ergaben zunächst, dass bei den religionskundlichen Grundkenntnissen 29% der Gesamtvarianz auf Schulunterschiede zurückzuführen ist, bei der hermeneutischen Fähigkeit sogar 41%. Kontrolliert man den Effekt der Schulform – in diesem Fall operationalisiert als Gymnasium vs. kein Gymnasium – so entfallen im Bereich der Kenntnisse 15%, im Bereich der hermeneutischen Fähigkeiten 23% der Gesamtvarianz auf Unterschiede zwischen dem Gymnasium und anderen Schulformen. Die Schulunterschiede berechtigen zu der Annahme, dass es auf der Ebene der Schule sowie der Schulform günstige und ungünstige Entwicklungsbedingungen für die religiöse Kompetenz gibt.

Die Varianzanteile¹⁹, die auch nach der Kontrolle von individuellen Lernvoraussetzungen der Schülerschaft (wie Geschlecht, familiären Buchbesitz, Familiensprache, religiöser Erfahrung sowie höchster Schulabschluss der Eltern) verbleiben, sind statistisch signifikant und praktisch bedeutsam, sodass die angenommene Abhängigkeit der religiösen Kompetenz von institutionellen, d.h. schulbezogenen und unterrichtsbezogenen Merkmalen nachweisbar ist. Dieser Befund deutet zunächst darauf hin, dass im Bereich der religiösen Kompetenz die schulform- und unterrichtsspezifischen Einflüsse von entscheidender Bedeutung sind. Die gefundenen Varianzanteile auf der Schul- und Schulformebene bestätigen darüber hinaus die Sensitivität für institutionelle Einflüsse der im Rahmen von RU-Bi-Qua/KERK entwickelten Skalen.

19 Nach der Kontrolle von individuellen Lernvoraussetzungen sowie der Schulformzugehörigkeit werden insgesamt 33% der Varianz der hermeneutischen Fähigkeiten erklärt; im Bereich der religionskundlichen Grundkenntnisse werden 32% der Gesamtvarianz erklärt. Von den Anteilen der Gesamtvarianz, die durch das vollständige Modell nicht erklärt werden konnten, sind für die Interpretationsskala 18% zwischen den Schulen und 81% zwischen den Schülern innerhalb der Schulen verteilt. Im Bereich der Kenntnisskala liegen 6% der nicht erklärten Varianzanteile auf Schulebene und 94% auf der Individualebene.

4. Fazit

Die religiösen Kompetenzen, gemessen nach dem Berliner Kompetenzmodell und dem in RU-Bi-Qua/KERK entwickelten Testinstrument, lassen sich einerseits als unterrichtsnah, d.h. durch das Curriculum im Religionsunterricht und dessen Umsetzung an der Einzelschule stark beeinflusst, beschreiben. Gleichwohl können auch die Effekte außerschulischer Determinanten mit dem Fragebogen erhoben und analysiert werden. Um den schulisch bedingten Unterschieden im Blick auf hermeneutische Fähigkeit und religionskundliche Grundkenntnisse weiter nachgehen zu können, ist allerdings ein längsschnittlich angelegtes oder ein quasi-experimentelles Untersuchungsdesign erforderlich. Dieses Ziel wird gegenwärtig in dem DFG-Projekt KERK angestrebt.

Für die weitere Arbeit im Projekt ergeben sich folgende Perspektiven:

In KERK werden zusätzlich zu den Aufgaben für die Testung von religiöser Deutungskompetenz Aufgaben entwickelt, mit denen religiöse Partizipationskompetenz getestet werden kann. Außerdem sollen der Schülerfragebogen optimiert und ein Lehrerfragebogen erarbeitet werden.

Zur externen Validierung der Ergebnisse wird in KERK ein Intelligenztest und ein Leseverständnistest eingesetzt werden. Die Projektgruppe strebt eine repräsentative Stichprobe von Zehntklässlern in Berlin und Brandenburg sowie, wenn dies gelingt, auch in anderen Bundesländern an.

Die dabei gewonnenen Ergebnisse sollen es ermöglichen, differenziert und empirisch valide religiöse Kompetenzniveaus zu beschreiben und einzuschätzen, welchen Einfluss erteilter Religionsunterricht an der öffentlichen Schule auf die Entwicklung religiöser Kompetenz als einem Bestandteil des Fähigkeitsspektrums der schulischen Allgemeinbildung hat.²⁰ Mithilfe einer Vergleichsgruppe soll ebenfalls untersucht werden, welche Folgen das Ausbleiben des Religionsunterrichtes für diesen Bereich der Allgemeinbildung zeigt.

Literatur

- Baumert, J. (2002): Deutschland im internationalen Bildungsvergleich. In: Killius, N. et al. (Hrsg.): Die Zukunft der Bildung. Frankfurt a.M.: Suhrkamp, S. 100–150.
- Benner, D. (2007): Unterricht – Wissen – Kompetenz. Deutungs- und Partizipationskompetenz im Religionsunterricht. Erscheint In: Schreiner, M. (Hrsg.): Gott Lesen, die Welt begreifen. Religious literacy. Münster: Waxmann.
- Benner, D./Krause, S./Nikolova, R./Pilger, T./Schluß, H./Schieder, R./Weiß, Th./Willems, J. (2007): Ein Modell domänenspezifischer religiöser Kompetenz. Erste Ergebnisse aus dem DFG-Projekt RU-Bi-Qua. In: Benner, D. (Hrsg.): Bildungsstandards. Instrumente zur Quali-

20 Nicht nur Jürgen Baumert schätzt die Probleme konstitutiver Rationalität, die er in den Fächern Religion und Philosophie verhandelt sieht (Baumert 2002, S. 113), sondern auch die PISA-Studie kennt neben „kognitiver, moralisch-evaluativer und ästhetisch-expressiver Rationalität“ einen Bereich „religiös-konstitutiver Rationalität“ (PISA 2001, S. 21), auch wenn er freilich in ihr nicht eigens untersucht wurde.

- tätssicherung im Bildungswesen. Chancen und Grenzen – Beispiele und Perspektiven. Paderborn: Schöningh, S. 141–156.
- Evangelische Kirche in Deutschland [EKD] (1994): Identität und Verständigung. Standort und Perspektiven des Religionsunterrichts in der Pluralität. Eine Denkschrift der Evangelischen Kirche in Deutschland. Gütersloh: Gütersloher Verlagshaus.
- Klieme, E./Avenarius, H./Blum, W. u.a. (2003): Expertise zur Entwicklung nationaler Bildungsstandards. Berlin: BMBF.
- Klieme, E./Funke, J./Leutner, D. u.a. (2001): Problemlösen als fächerübergreifende Kompetenz. Konzeption und erste Resultate aus einer Schulleistungsstudie. In: Zeitschrift für Pädagogik 47, S. 179–200.
- Matthes, J. (1992): Auf der Suche nach dem Religiösen. Reflexion zu Theorie und Empirie religionssoziologischer Forschung. In: Sociologia Internationalis 30, S. 129–142.
- Nikolova, R./Schluss, H./Weiß, Th./Willems, J. (2007): Das Berliner Modell religiöser Kompetenz. Fachspezifisch-Testbar-Anschlußfähig. In: www.theo-web.de, 2/2007.
- PISA-Konsortium (Hrsg.) (2001): PISA 2001. Opladen: Leske & Budrich.
- Schieder, R. (2007): Modell 3: Verordnete, gefühlte oder messbare Bildungsstandards? Konzeption und Forschungsstand eines interdisziplinären Berliner Projekts. In: Sajak, C.P. (Hrsg.): Bildungsstandards für den Religionsunterricht – und nun? Perspektiven für ein neues Instrument im Religionsunterricht. Münster u.a.: LIT, S. 67–86.

Abstract: *As is the case in all special didactics, in the field of religious instruction, too, there is a tendency to describe subject-related competence in such a way that it becomes amenable to empirical analysis. The authors present the concept of the DFG-sponsored projects "Quality Maintenance and Educational Standards in Religious Instruction at State Schools as exemplified by Protestant Religious Instruction" and "Construction and Survey Of Religious Competence Levels in Religious Instruction as exemplified by Protestant Religious Instruction" (RU-Bi-Qua/KERK). In these projects, educational scientists and teachers of religion worked together to develop a concept of religious competence as well as questions for its empirical testing. Here, the results of this construct validation study are discussed for the first time. They confirm the approach chosen at crucial points, however, they also show the direction future research should take.*

Adressen der Autorinnen und Autoren:

Sabine Krause/Roumiana Nikolova/Dr. Henning Schluß, Humboldt-Universität zu Berlin, Institut für Erziehungswissenschaften, Unter den Linden 6, 10099 Berlin,
E-Mail: sabine.kra@web.de; roumiana.nikolova@rz.hu-berlin.de;
henning.schluss@rz.hu-berlin.de
Dr. Thomas Weiß/Dr. Joachim Willems, Humboldt-Universität zu Berlin, Seminar für Praktische Theologie/Religionspädagogik, Unter den Linden 6, 10099 Berlin,
E-Mail: dr.tom.weiss@gmx.net; joachim.willems@gmx.de

Detlev Leutner

Metamorphose eines Forschungsprojektes

Ein Kommentar zum Beitrag von Krause et al. über den „Kompetenzerwerb im evangelischen Religionsunterricht – Ergebnisse der Konstruktvalidierungsstudie der DFG-Projekte RU-Bi-Qua / KERK“

Zusammenfassung: *Der hier kommentierte Artikel beschreibt ein Projekt, das im Laufe seiner Entwicklung eine Art Metamorphose durchlaufen hat: Das Projekt begann mit einer weit gefassten, normativ-anwendungsorientierten Fragestellung und betreibt nun, nach Aufgreifen konstruktiv-kritischer Anregungen aus einem zweistufigen „Peer Review“-Verfahren, empirisch orientierte, anschlussfähige Grundlagenforschung zur Kompetenzmodellierung auf methodisch hohem Niveau in einem fachlichen Kontext, in dem derartige Forschung bisher wenig üblich war.*

1. Einleitung

Die von Krause et al. (2008) vorgestellten Forschungsergebnisse entstammen einem erfolgreichen, zunächst mit einer Laufzeit von 18 Monaten geförderten und zwischenzeitlich um weitere zwei Jahre verlängerten Drittmittelprojekt, das ursprünglich die Formulierung von Bildungsstandards für den evangelischen Religionsunterricht zum Ziel hatte, in der Verlängerung nun aber auf die Konstruktion und die Erhebung von religiösen Kompetenzniveaus im Religionsunterricht fokussiert. Dieser Wandel in der Zielsetzung des Projektes lässt sich als Beispiel dafür beschreiben, wie aus einem Projekt mit einer vom Ansatz her eher normativ-anwendungsorientierten Fragestellung Grundlagenforschung entstehen kann, die nicht nur theoretisch und praktisch relevante Ergebnisse zu erbringen vermag, sondern darüber hinaus auch in ihrem empirisch-methodischen Vorgehen Ansprüchen gerecht wird, die in anderen Bereichen längst „State-of-the-Art“ sind. Die Metamorphose dieses Projektes soll im Folgenden vor dem Hintergrund gängiger Standards grundlagenorientierter empirischer Bildungsforschung zur Kompetenzmodellierung kommentiert werden.

2. Von einer normativ-anwendungsorientierten Fragestellung zu empirisch orientierter Grundlagenforschung

2.1 Der ursprüngliche Projektplan

Zielsetzung: Wie einleitend schon beschrieben, bestand das ursprüngliche Ziel des Projektes darin, Bildungsstandards für den evangelischen Religionsunterricht zu formulieren. Die Setzung von Standards, so auch die Setzung von Bildungsstandards (vgl. Klieme et al. 2003), ist eine Aufgabe, die typischerweise von staatlichen Institutionen (in