

Wirksamkeitsforschung in der Religionsdidaktik: Ziele, Forschungsfelder und die Frage nach religionsdidaktischen Forschungsformaten

von

Friedrich Schweitzer

Dieser Beitrag beschreibt Ansätze der auf den Religionsunterricht bezogenen Wirksamkeitsforschung in religionsdidaktischer Perspektive. Beschrieben werden mögliche Ziele solcher Ansätze sowie bislang fassbare Forschungsfelder. Darüber hinaus wird der Begriff verschiedener Forschungsformate aufgenommen, der in der neueren deutschen Diskussion über Allgemeine Fachdidaktik eingeführt worden ist. Dieser Begriff wird im Verhältnis zu auf den Religionsunterricht bezogener Wirksamkeitsforschung erörtert, um so seinen Nutzen zu prüfen und seinen religionspädagogischen Sinn zu klären.

This article describes approaches to investigating the effects of RE in the perspective of religion-related didactics. It describes possible aims of such approaches as well as existing fields of research. At the same time, the article stakes up the concept of different research formats which has been introduced by the recent German discussion on general subject-oriented didactics. This concept is discussed in relationship to research on the effects of RE, thus testing the possible usefulness of this concept and trying to clarify its meaning for religious education.

Schlagworte: Wirksamkeitsforschung; Religionsdidaktik; Forschungsformate; empirische Ansätze; Methodologie

Keywords: effect-oriented research; religious education; research formats; empirical approaches; methodology

Wer Unterricht veranstaltet oder selbst unterrichtet, will etwas bewirken. Unterrichten ist eine zielgerichtete Aktivität, die sich darin von anderen Tätigkeiten wie etwa gemeinsamem Spiel, Unterhaltung usw. unterscheidet. Insofern ist die Frage nach der Wirksamkeit der Unterrichtspraxis selbst immanent, und noch mehr gilt dies für die Wissenschaft, die sich als (Fach-)Didaktik um die wissenschaftliche Erforschung und Weiterentwicklung dieser Praxis bemüht. Auch sie kann auf die Frage nach der Wirksamkeit nicht gut verzichten, auch wenn die (Fach-)Didaktik dies in der Vergangenheit über weite Strecken getan oder die Wirksamkeitsfrage nur nebenbei beachtet hat (zum weiteren Hintergrund vgl. Schweitzer, 2020).

Zugleich geht Unterricht nicht in seiner Wirksamkeit auf, vor allem nicht im Erwerb von Kompetenzen, die als Outcome verstanden werden und die sich oft erst später nutzen lassen. Die Pädagogik hat zu Recht darauf aufmerksam gemacht, dass eine solche Verzweckung des Lernens oder gar des kindlichen Lebens für die Zukunft nicht legitim wäre. Friedrich Schleiermacher hat dies in seiner „Erziehungslehre“ anthropologisch und ethisch begründet (Schleiermacher, 1849, S. 70–79). Insofern gilt, dass der Weg auch das Ziel sein muss. Beim Unterricht kommt es nie allein auf außerhalb des Unterrichts fassbare Wirkungen im Sinne externer Zwecke an, sondern immer auch auf das, was direkt im Unterricht erfahren werden kann.

Die Überzeugung, dass diese Sicht besonders für den Religionsunterricht gilt, erklärt vermutlich zumindest teilweise, warum es die Wirksamkeitsforschung hier – etwa im

Unterschied zu den naturwissenschaftlichen Fächern – besonders schwer gehabt hat. Dazu kommen weitere Grenzen, die aus den Themen des Religionsunterrichts resultieren: Besonders wenn es um existenzielle Fragen oder den Glauben geht, kann der Religionsunterricht zwar faktisch durchaus Wirkungen erzeugen, aber es wäre weder pädagogisch noch theologisch legitim, Glauben bewirken zu wollen, wie insbesondere in der evangelischen Tradition („Nicht-Lehrbarkeit des Glaubens“) immer wieder betont wurde (Schweitzer, 2019, S. 26–37). Ähnliches gilt auch für die personale Dimension von Unterricht in der Gestalt von Beziehungen zwischen Religionslehrkräften und Schülerinnen und Schülern, die sich nicht sinnvoll in eine wirkungsbezogene Terminologie im hier gemeinten Sinne übersetzen lässt, auch wenn solche Beziehungen natürlich Auswirkungen haben können. An allen diesen Punkten wird zugleich deutlich, dass zwischen *Wirkungen* in einem *weiteren* und in einem *engeren Sinne* unterschieden werden muss. Wirkungen im weiteren Sinne betreffen das Leben insgesamt und sind entsprechend schwer zu fassen. Wirkungen von Unterricht im engeren Sinne hingegen bezeichnen konkrete, insofern definierbare und häufig empirisch fassbare Lernergebnisse, die nach einer Unterrichtsstunde, einer Unterrichtseinheit oder eines Schuljahrs nachweisbar sein müssen. Es wäre wenig ratsam, wollte sich der Religionsunterricht allein auf kaum nachprüfbar, erst in einer unbestimmten Zukunft greifbare Wirkungen, berufen.

Vor diesem Hintergrund lässt sich die *erste Aufgabe* des vorliegenden Beitrags bestimmen: Geklärt werden muss, welche Ziele genau die Wirksamkeitsforschung in der Religionsdidaktik verfolgen soll und wie dabei vorzugehen ist. Mit der vorliegenden Publikation verbindet sich aber noch eine *zweite Aufgabe*, die weiterreichend das Verständnis von Forschung in der Religionsdidaktik betrifft: Es soll geprüft werden, ob sich die religionsdidaktische Wirksamkeitsforschung im Sinne eines religionsdidaktischen *Forschungsformats* verstehen lässt. Damit verbinden sich verschiedene interessante Fragen: Zunächst ist genauer zu klären, was der Formatbegriff im vorliegenden Zusammenhang bedeuten kann oder bedeuten soll. Sodann ist zu untersuchen, welchen Gewinn es verspricht, wenn der aus der aktuellen Diskussion in der fachdidaktischen Forschung stammende Formatbegriff in die Religionsdidaktik übernommen werden soll. Damit bezieht sich der vorliegende Beitrag insbesondere auf Fragen im Umkreis der Allgemeinen Fachdidaktik, die sich auch auf Forschungsaufgaben der Didaktik erstrecken (Bayrhuber, Harms, Muszynski, Ralle, Rothgangel, Schön, Vollmer & Weigand, 2012; Rothgangel, Abraham, Bayrhuber, Frederking, Jank & Vollmer, 2020). Allerdings ist bislang auch in diesem Bereich noch wenig geklärt, was der Formatbegriff dabei genau bedeutet.

1 Ziele religionsdidaktischer Wirkungsforschung

Religionsdidaktische Wirksamkeitsforschung fragt nach Wirkungen von Religionsunterricht, weshalb zunächst das religionspädagogische Verständnis von Wirkung geklärt werden muss. Hilfreich ist dabei die bereits erwähnte Unterscheidung zwischen Wirkungen im engeren und im weiteren Sinne. Im engeren Sinne geht es demnach um die Frage, ob die im Religionsunterricht im Blick auf die Schülerinnen und Schüler bzw. deren Erwerb von Kenntnissen, Fähigkeiten, Fertigkeiten usw. verfolgten Ziele nach der Durchführung des Unterrichts in nachweisbarer Form tatsächlich erreicht sind; bei Wirkungen im weiteren Sinne hingegen ist auch an Effekte zu denken, die wie etwa die allgemeine Resonanz des Faches im Sinne seiner Beliebtheit betreffen, die aber nicht zum üblichen Planungshorizont für eine bestimmte Stunde, Unterrichtseinheit oder

auch einen längeren Zeitraum im Schuljahr gehören. Dazu zählen darüber hinaus Effekte wie ein sich vielleicht langfristig – positiv oder negativ – einprägendes Lehrer-Schüler-Verhältnis usw. Im Folgenden soll die Erforschung von Wirksamkeit im engeren Sinne im Zentrum stehen.

Bereits einleitend wurde ebenfalls darauf aufmerksam gemacht, dass der Wirksamkeitsforschung im Religionsunterricht *prinzipielle Grenzen* gezogen sind, die aus der Thematik dieses Unterrichts erwachsen: Wo immer es um *existenzielle Fragen* oder um den *Glauben* geht, ist eine solche Wirksamkeitsforschung pädagogisch und theologisch als illegitim abzulehnen. Da es sich dabei um eine wesentliche Dimension von Religionsunterricht handelt, ist diese Einschränkung durchaus bedeutsam. Diese Einschränkung darf aber gerade deshalb nicht als ein grundlegender Einwand gegen jede Form von Wirksamkeitsforschung ausgegeben werden. Auch dies wäre weder pädagogisch noch theologisch begründbar.

Ein Unterricht, in dem keine Kenntnisse, Fähigkeiten und Fertigkeiten erworben werden, die sich auch überprüfen lassen, hätte in der Schule schwerlich einen Platz. Er würde aufhören, Unterricht im Sinne der Schule zu sein. Das gilt auch für den Religionsunterricht, der eben immer beides ist: ein ganz besonderes Fach, aber eben doch zugleich ein Fach wie jedes andere.

Zur Überprüfung der Wirksamkeit von Unterricht wird heute der Erwerb von Kenntnissen, Fähigkeiten und Fertigkeiten mithilfe von Kompetenzmessungen erfasst. Dies bedeutet ausdrücklich nicht, dass deshalb in einem bestimmten Sinne etwa „kompetenzorientiert unterrichtet“ werden müsste (so noch etwa Obst, 2008). Vielmehr geht es um Kompetenzmodelle, die sich für jede Form von Unterricht einsetzen lassen. Denn Kenntnisse, Fähigkeiten und Fertigkeiten oder eben Kompetenzen sollen ja in jedem Unterricht erworben werden.

In der Religionsdidaktik gibt es bekanntlich keinen Konsens über ein Kompetenzmodell (vgl. zuletzt Riegel & Rothgangel, 2020a), und von einem auch empirisch validierten Kompetenzmodell, das diesem Ansatz zufolge eine entscheidende Voraussetzung darstellt (etwa nach Klieme et al., 2007), kann nur in Ansätzen gesprochen werden. Die breiteste Zustimmung hat in der Religionspädagogik bislang wohl das Modell der Berliner Forschergruppe um Dietrich Benner gefunden, das drei Teilkompetenzen oder Komponenten einschließt: Wissen, Deutungs- und Handlungsfähigkeit (Benner, Schieder, Schluß & Willems, 2011). Genauer gesagt bezieht sich die Zustimmung nur auf die ersten beiden Komponenten, während sich die Handlungsfähigkeit empirisch nicht abbilden ließ. In dieser Hinsicht stößt die Schule wohl auch an prinzipielle Grenzen, weil eine solche Kompetenz innerhalb der Schule kaum nachgewiesen werden kann. Es geht dabei ja gerade um ein Handeln außerhalb von Unterricht und Schule.

In neuerer Zeit wurden aufbauend auf dem Modell von Benner sowohl theoretische als auch forschungsbezogene empirische Differenzierungen eingeführt. Besonders wichtig ist dabei die Komponente der Perspektivenübernahmefähigkeit, die in der heutigen Religionspädagogik, die sich auf kulturelle und religiös plurale Verhältnisse einstellen muss, entsprechend gewichtet wird (vgl. exemplarisch Schweitzer, Bräuer & Boschki, 2017; Meyer, 2019). Dabei ist in dieser Hinsicht empirisch gesehen noch weithin offen, ob sich die Perspektivenübernahme als Teil der Deutungs-fähigkeit verstehen lässt – so noch bei Benner et al. (2011, S. 126) – oder ob sie als eine spezifische Form von Deuten oder Verstehen aufzufassen ist, auf die es in interreligiösen Zusammenhängen ankommt (so Schweitzer et al., 2017). Zugleich stellt mitunter bereits die empirische Erfassung von Perspektivenübernahmefähigkeit vor ungelöste Probleme (Schweitzer & Bucher, 2020). Ehe solche Fragen auch nur ein Stück weit abschließend beantwortet

werden können, müssen vor allem größere, möglichst repräsentative Studien durchgeführt werden, die sich zur Validierung von Kompetenzmodellen eignen. Solche Studien sind in der Religionspädagogik bislang nur für einzelne Bereiche durchgeführt worden, vor allem zum interreligiösen Lernen (dazu noch unten).

Grundsätzlich geht es jedoch nicht allein um die begriffliche Erfassung und Validierung eines Kompetenzmodells für den Religionsunterricht. Weiterreichend ist vielmehr der *Kompetenzerwerb* entscheidend, denn darin liegt ja der Sinn von Unterricht. Wissenschaftlich gesehen handelt es sich beim Kompetenzerwerb um eine Veränderung in der Zeit – um ein Vorher und Nachher, also beispielsweise um den Wissensstand vor einer Unterrichtseinheit und nach dieser Einheit. Forschungsdesigns, die solche Veränderungen verlässlich erfassen sollen, arbeiten deshalb in der Regel mit zwei oder mehr Befragungs- bzw. Messzeitpunkten.

Noch allgemeiner formuliert lässt sich sagen, dass die Wirksamkeitsforschung in der Religionsdidaktik immer zur Verbesserung von Religionsunterricht dienen soll und auf dieses Ziel eingestellt sein muss (Schweitzer, 2020). Deshalb sollte auch nicht der Aspekt von Evaluation im Sinne einer Bewertung von Unterricht im Vordergrund stehen, sondern vielmehr die Eröffnung von Entwicklungsmöglichkeiten. Auch solche Überlegungen müssen gerade für fachdidaktische Forschung von Anfang an beachtet werden.

Ehe genauer auf die verschiedenen Zugänge in der religionspädagogischen Wirksamkeitsforschung eingegangen wird, soll zunächst eine Übersicht über zentrale Forschungsfelder und -themen in diesem Bereich religionsdidaktischer Forschung versucht werden.

2 Zentrale Forschungsfelder religionsdidaktischer Wirksamkeitsforschung

Die erste Studie in der Religionsdidaktik, die im engeren Sinne als wirksamkeitsbezogen bezeichnet werden kann, wurde vor mehr als 30 Jahren von Fritz Oser (1988) durchgeführt. Im Zentrum stand hier die Frage, ob sich durch entsprechenden („stimulierenden“) Unterricht die Entwicklung des religiösen Urteils erreichen lässt. Darin folgte Oser dem Vorbild von Lawrence Kohlberg (1995), der schon früher entsprechende Untersuchungen zur Entwicklung des moralischen Urteils auch im Verhältnis zu Unterricht durchgeführt hatte. Auch wenn es sich bei Kohlberg um Untersuchungen und nicht um Studien zum Religionsunterricht handelte, geht es auch bei ihm um einen Themenbereich, der für den Religionsunterricht besonders relevant ist.

Für die Gegenwart legt sich die Unterscheidung zwischen Wirksamkeitsstudien zum *Religionsunterricht als solchem* auf der einen und Untersuchungen zu *verschiedenen Formen der Ausgestaltung von Religionsunterricht*, etwa im Sinne spezieller Unterrichtseinheiten oder didaktischer Ansätze, nahe (Überblick: Schweitzer & Boschki, 2018). Bei den Untersuchungen zum Religionsunterricht als solchem wird gefragt, welche Veränderungen sich bei Kindern oder Jugendlichen, die den Religionsunterricht besuchen, im Vergleich zu anderen, die diesen Unterricht nicht besuchen, feststellen lassen. Ein dafür besonders geeignetes Untersuchungsfeld bietet in Deutschland das Land Berlin, in dem der Religionsunterricht nur als freiwilliges Angebot existiert und daher auch nicht von allen Schülerinnen und Schülern besucht wird. Bei diesem Vergleich konnte die bereits erwähnte Forschungsgruppe um Benner einen deutlichen Vorsprung an Wissen und Deutungsfähigkeit für diejenigen nachweisen, die am Religionsunterricht teilgenommen hatten (Benner et al. 2011, S. 135–140). In einer späteren Untersuchung zum Ethikunterricht bzw. zu Praktischer Philosophie in Nordrhein-

Westfalen konnten ähnliche Befunde erzielt werden, wobei sich der Philosophieunterricht im Blick auf ethische Kompetenzbildung als etwas wirksamer erwies als der Religionsunterricht, vor allem aber die an einem der beiden Angebote beteiligten Schülerinnen und Schüler deutlich höhere Kompetenzausprägungen aufwiesen als diejenigen, bei denen keine Teilnahme gegeben war (Ivanov, 2016, S. 187–194). In Salzburg konnte Georg Ritzer (2010) mit einer Untersuchung zum Besuch von Religionsunterricht während eines Schuljahrs ebenfalls zeigen, dass besonders im Blick auf religionsbezogenes Wissen deutliche Zuwächse festzustellen waren. Der Nachteil solcher Untersuchungen, die einfach nach der Wirksamkeit der Teilnahme am Religionsunterricht fragen, bzw. ihren Grenzen, besteht darin, dass der besuchte Unterricht nicht genauer erfasst wurde. Es ging lediglich um Teilnahme oder Nicht-Teilnahme. Im internationalen Bereich ist hier auf die Studie von James Conroy zu verweisen, die ebenfalls nach der Wirksamkeit von Religionsunterricht fragte und dabei eine zwar hoch komplexe, zugleich jedoch nicht leicht nachvollziehbare Kombination mehrerer Untersuchungsmethoden einsetzte und zu eher skeptischen Einschätzungen im Blick auf die Wirksamkeit des Unterrichts gelangte (Conroy, Lundie, Davis, Baumfield, Barnes, Gallagher, Lowden, Bourque & Wenel, 2013). Andere Studien aus dem United Kingdom, die im Anschluss daran durchgeführt wurden, gelangen etwa hinsichtlich des Bildungsziels der Pluralitätsfähigkeit zu positiveren Befunden (Francis, Penny & McKenna, 2016).

Um gezieltere Befunde erreichen zu können, bei denen die jeweilige Ausgestaltung des Religionsunterrichts und damit auch religionsdidaktische Fragestellungen berücksichtigt werden können, bieten sich in der Religionsdidaktik ähnlich wie in anderen Fächern oder Feldern des Lernens sogenannte Interventionsstudien an, bei denen eine bestimmte Form der Gestaltung von Religionsunterricht, etwa in Gestalt einer Unterrichtseinheit oder eines didaktischen Ansatzes wie der Schülerorientierung, auf den Prüfstand gestellt wird (Schweitzer, 2018). Kennzeichnend für Interventionsstudien ist dabei der Einbezug einer Vergleichsgruppe, die nicht an dem entsprechenden Angebot teilnimmt. Darüber hinaus werden die Eingangsvoraussetzungen vor dem Beginn der Unterrichtseinheit erfasst, und eine zweite Messung erfolgt am Ende der Unterrichtseinheit. Um auch die Nachhaltigkeit des Lernens in den Blick nehmen zu können, wird darüber hinaus häufig im zeitlichen Abstand zum Unterricht eine dritte Messung durchgeführt. Solche Untersuchungen sind naturgemäß sehr aufwändig, was erklärt, warum im Religionsunterricht bislang erst eine überschaubare Anzahl von Interventionsstudien verfügbar geworden ist. Themenbereiche solcher Studien waren insbesondere interreligiöses Lernen (s.u.), Wertebildung (s. ebenfalls unten) oder das Schöpfungsverständnis (Göbinger, 2014; Rodegro, 2010). In allen diesen Fällen waren durchaus Lernerfolge zu konstatieren, aber es ergaben sich auch weitreichende Rückfragen im Blick auf die Zielsetzung der entsprechenden didaktischen Ansätze. Um nur einige zu nennen:

- Die in der religionsdidaktischen Literatur vielfach genannte Fähigkeit zur *Perspektivenübernahme* erweist sich als empirisch nur schwer fassbar, und entsprechend fällt es schwer, hier Kompetenzzuwächse verlässlich zu überprüfen (Schweitzer & Bucher, 2020; zur weiteren Diskussion s. auch Losert, Merkt & Schweitzer, 2015). Diese Schwierigkeit betrifft dabei nicht nur die Forschung, sondern entsprechend auch die Unterrichtspraxis. Denn wenn es in der Forschung nicht gelingt, Kompetenzzuwächse bei der Fähigkeit zur Perspektivenübernahme nachzuweisen, dann kann dies auch für die Unterrichtspraxis kaum vorausgesetzt werden.

- Besonders herausfordernd ist der sich wiederholende Befund, dass Änderungen bei *Einstellungen* im untersuchten Religionsunterricht nicht festzustellen waren (Ziebertz, 2010; Ritzer, 2010; Schweitzer et al., 2017). Dieser Befund ist deshalb so herausfordernd, weil entsprechende Effekte vom Religionsunterricht gerade auch in Politik und Öffentlichkeit erwartet werden, und ebenso in der Religionsdidaktik selbst als sehr wünschenswert angesehen werden – man denke etwa an den neuen Antisemitismus oder die Islamfeindlichkeit in Deutschland. Worin das Ausbleiben entsprechender Effekte begründet ist, muss noch genauer untersucht werden. Mitunter wird erwartet, dass die bei einer Intervention mögliche Zeitspanne von vielleicht wenigen Wochen einfach zu klein sei, um mit solchen Effekten überhaupt rechnen zu können. Dieser Einwand klingt durchaus plausibel. Gleichwohl bleibt die Rückfrage, ob sich nicht auch bei einer kurzen Zeitspanne zumindest kleine Effekte abzeichnen müssten, aus denen sich dann kumulativ größere Effekte ergeben könnten.
- Soweit zu den Einstellungen auch *Interesse* zählt, kann hier auf gewisse Erfolge von Religionsunterricht verwiesen werden (im Bereich der Wertebildung: Wagen-sommer & Schweitzer, 2018). Es kann im Religionsunterricht offenbar gelingen, Interesse zu steigern. Allerdings ist die empirische Grundlage für solche Aussagen noch überaus schmal und andere Studien zeigen, dass dies im Religionsunterricht keineswegs immer erreicht wird.
- Im Blick auf die Religionsdidaktik besonders erwähnenswert ist der Befund aus einer aufwändigen Interventionsstudie zum interreligiösen Lernen, bei der zugleich subjekt- bzw. lebensweltorientierte didaktische Ansätze eine besondere Rolle spielten (Schweitzer et al., 2017). Hier zeigte sich, dass es gerade solche didaktischen Vorgehensweisen waren, die nicht in hervorgehobener Weise subjekt- oder lebensweltorientiert waren, die eine größere Wirksamkeit entfalteten. Dies widerspricht natürlich der religionsdidaktischen Tradition und sollte deshalb dringend weiter untersucht werden. Eine Hypothese geht hier dahin, dass die didaktische Rede von Subjekt- bzw. Lebensweltorientierung präzisiert werden muss.

Religionsdidaktik ist nicht auf den schulischen Religionsunterricht beschränkt, sondern kann und sollte auch andere Lern- und Organisationsformen im Raum der Schule einbeziehen. Das zeigt sich schon an zwei Studien, die auf die Wirksamkeit von Besuchen außerschulischer Lernorte gerichtet waren, zum einen der Besuch unterschiedlicher jüdischer bzw. auf das Judentum bezogener Einrichtungen, zum anderen von Kirchen im Sinne der Kirchenpädagogik (Gärtner & Bettin, 2015; Riegel & Kindermann, 2017). In beiden Fällen konnte die Attraktivität der entsprechenden Lernform nachgewiesen werden, beispielsweise anhand von Wirkungen auf das Interesse am Thema. Zugleich ergaben sich Unterschiede in der Wirksamkeit, abhängig von der jeweiligen Ausgestaltung der Besuche sowie der Auswahl der besuchten Orte. Dieser Befund verweist darauf, dass es nicht einfach um eine bestimmte Lernform als solche geht, auf die hinsichtlich der Wirksamkeit von Unterricht zu achten ist. Ähnlich wie bei der Subjektorientierung von Religionsunterricht kommt es offenbar auch in diesem Falle darauf an, in welcher Weise tatsächlich gearbeitet wird.

Auch die Wirksamkeit non-formalen religionsbezogenen Lernens, das nicht mit der Schule verbunden ist, kann untersucht werden (Überblick: Schweitzer, Ilg & Schreiner, 2019). Am weitesten reichen dabei bislang die Untersuchungen zur Konfirmandenarbeit (Überblick: Simojoki, Ilg, Schlag & Schweitzer, 2018), bei der sich zwar (noch) nicht im Sinne von Interventionsstudien Wirkungszusammenhänge untersuchen ließen, wohl aber mithilfe etwa von Regressions- und Mehrebenenanalysen Prädiktoren

für bestimmte erwünschte Effekte identifiziert werden konnten. Beispielsweise ergab sich ein Zusammenhang zwischen der Teilnahme an „Konfi-Camps“ und der Gesamtzufriedenheit mit der Konfirmandenarbeit oder zwischen der Möglichkeit, während der Konfi-Zeit an zeitlich begrenzten Praktika teilzunehmen, und einem späteren ehrenamtlichen Engagement bzw. der Bereitschaft dazu (Ilg, Pohlers, Gräbs Santiago & Schweitzer, 2018).

In neuerer Zeit wird auch nach der Wirksamkeit von Aus- und Fortbildungsangeboten für Lehrkräfte gefragt (Rzejak & Lipowsky, 2019), wobei hier im Bereich des Religionsunterrichts erst von Anfängen einer solchen Wirksamkeitsforschung gesprochen werden kann. Auch Fortbildungsangebote lassen sich im Sinne von Interventionen verstehen, aber es ist nicht leicht, die möglichen Effekte von Fortbildung auf die Unterrichtsgestaltung verlässlich zu erfassen. Ähnlich schwer zu fassen ist der Zusammenhang zwischen Ausbildung und späterer Unterrichtspraxis. Gleichwohl ist es zu begrüßen, wenn auch dieser Bereich der Religionsdidaktik nicht aus dem Blick gerät.

3 Forschungsdesigns und Methoden

Wie bereits deutlich geworden ist, sind für die Wirksamkeitsforschung zunächst die gewählten Forschungsdesigns entscheidend, nicht einfach die Methoden. In der empirischen Bildungsforschung figurieren allerdings auch die Designs mitunter unter dem Begriff der Methodik.

Ein typisches Design im Bereich der Wirksamkeitsforschung dieser Art sind die beschriebenen Interventionsstudien. Daneben wird aber nach wie vor auch mit weniger aufwändigen Designs gearbeitet – vor allem mit Regressionsanalysen, die allerdings nur in bestimmten Hinsichten Aufschlüsse bieten können. In der Regel finden sie ihre Grenze darin, dass der jeweils erfahrene Unterricht in seiner jeweiligen Ausgestaltung dabei nur indirekt erfasst wird (bspw. Francis et al., 2018).

Die lange Zeit als entscheidend auch für die Unterrichtsforschung angesehene Unterscheidung zwischen *quantitativen* und *qualitativen* Zugangsweisen ist im Blick auf die religionsdidaktische Wirksamkeitsforschung nur bedingt hilfreich. Veränderungen lassen sich sowohl mithilfe qualitativer als auch quantitativer Zugangsweisen erfassen, aber verlässliche Vergleiche allein auf qualitativer Basis sind überaus aufwändig und lassen in der Regel eine repräsentative Grundlage vermissen. In der quantitativen Forschung mit Interventionen zeichnet sich demgegenüber ein Trend zu sogenannten *Randomized Controlled Trials* ab, d.h. es werden Anforderungen an *zufallsgenerierte Samples* in der Praxis erhoben, wie sie in anderen Wissenschaftsbereichen wie etwa der Medizin ebenfalls üblich sind und als „gold standard“ bezeichnet werden (Schulz, Altman & Moher, 2010). Daran ist abzulesen, wie wichtig die Kriterien für die Anlage und Zusammensetzung von Untersuchungsgruppen eingeschätzt werden. Darüber hinaus wachsen zugleich die *Ansprüche an die Auswertungsformen*, bei denen Effektstärken, Item-Response-Analysen, Mehrebenenanalysen usw. zunehmend zum Standard geworden sind. Solche Verfahren sind in der Religionsdidaktik bislang noch sehr selten zur Anwendung gekommen. In der Praxis setzen sie eine Expertise voraus, die in der Religionsdidaktik selbst kaum einmal anzutreffen ist.

Untersuchungen, bei denen ausschließlich quantitative Verfahren zum Einsatz kommen, sind in der religionsdidaktischen Wirksamkeitsforschung inzwischen eher selten. Zumeist werden bei quantitativen Untersuchungen qualitative Vorstudien oder Vertiefungsstudien durchgeführt, wofür dann auch Begriffe wie Mixed-Methods-Ansatz stehen. Bei vorwiegend qualitativen Untersuchungen zur Wirksamkeit von Unterricht

kommt es mitunter am Ende doch zu einer gewissen Quantifizierung der Befunde, eben weil die Frage nach Wirksamkeit exakte Vergleiche voraussetzt.

Eine interessante Frage betrifft das Verhältnis zwischen Wirksamkeitsuntersuchungen zum Religionsunterricht und der religionsdidaktischen Entwicklungsforschung. Bei Riegel & Rothgangel (2020b) werden sie als unterschiedliche Forschungsformate geführt. Diese Unterscheidung darf aber nicht darüber hinwegtäuschen, dass sich diese beiden Ansätze notwendigerweise überlappen. Denn bei der religionsdidaktischen Entwicklungsforschung kommt es ja darauf an, dass die jeweils entwickelten didaktischen Vorgehensweisen in der Praxis erprobt, evaluiert und dann verbessert werden (Gärtner, 2018a). Wie aber soll eine evaluative Erprobung aussehen, wenn dabei nicht nach der Wirksamkeit gefragt wird? Gärtner (2018b, S. 26) und Prediger (2018, S. 172) sind allerdings skeptisch, was die Durchführbarkeit entsprechender Wirksamkeitsforschung im Rahmen der religionsdidaktischen Entwicklungsforschung betrifft. Die religionsdidaktische Entwicklungsforschung erscheint ihnen dafür noch zu wenig ausgebildet, während zugleich die entsprechenden Effekte zu komplex für eine verlässliche Erfassung von Wirkungen seien. Es ist jedoch nicht einsichtig, dass in solchen pragmatischen Erwägungen ein prinzipielles Argument gesehen werden soll. Vielmehr ist hier so zu urteilen, dass der bisherige Stand der religionsdidaktischen – und vielleicht auch sonst der fachdidaktischen – Entwicklungsforschung bei der Evaluation von unterrichtlichen Erprobungen nur pragmatische Einschätzungen zulässt, obwohl valide Untersuchungen zur Wirksamkeit erforderlich wären. Insofern kommt es dann entscheidend darauf an, in Zukunft geeignete Instrumente und Vorgehensweisen zu entwickeln, die es erlauben, über die pragmatische zu einer systematischen Vorgehensweise zu gelangen.

Die derzeit vielleicht wichtigste einschränkende Voraussetzung für die religionsdidaktische Wirksamkeitsforschung kann bei alle dem darin gesehen werden, dass auf den Religionsunterricht bezogene Untersuchungen, wie sie in den letzten 20 Jahren vermehrt verfügbar geworden sind, nur wenig Konstanz im Blick auf Fragestellungen und Instrumente aufweisen (so kritisch auch Schwarz, 2019). Auf diese Weise fällt es überaus schwer oder ist es überhaupt unmöglich, dass am Ende ein kumulativer Erkenntnisfortschritt erzielt wird. Denn unterschiedlich angelegte Studien führen nicht zu vergleichbaren Ergebnissen. Auch Aufschlüsse über, in diesem Bereich zu erwartende oder nicht zu erwartende, Wirkungen und deren Stärke lassen sich unter diesen Voraussetzungen kaum gewinnen. Darüber hinaus ist festzustellen, dass in vielen Untersuchungen vor allem nach der persönlich empfundenen Relevanz oder Attraktivität von Religionsunterricht gefragt wird. Das Unterrichtsgeschehen im engeren Sinne kommt vergleichsweise wenig in den Blick, selbst wenn immer wieder nach dem „Binnengeschehen“ im Unterricht gefragt wird. In der empirischen Bildungsforschung gibt es dazu weiterführende Angebote von Instrumenten, die in Zukunft auch stärker für den Religionsunterricht genutzt werden könnten und genutzt werden sollten (Göllner, Wagner, Klieme, Lüdtke, Nagengast & Trautwein, 2016).

4 Herausforderungen und Anfragen

Nachdem bereits am Ende des letzten Abschnitts auf einige Probleme verwiesen wurde, soll an dieser Stelle vor allem auf zwei Herausforderungen bzw. Anfragen aufmerksam gemacht werden. Die erste ist eher pragmatischer Natur, die zweite wirft grundsätzliche Fragen auf.

Besonders Interventionsstudien, die am verlässlichsten über mögliche Wirkungen Auskunft geben können, sind in der Durchführung überaus anspruchsvoll. Ohne eine starke Drittmittelfinanzierung lassen sie sich kaum realisieren. Darüber hinaus ist – wie sich gezeigt hat – eine sozialwissenschaftlich-statistische Expertise erforderlich, wie sie in der Religionspädagogik selbst in aller Regel nicht vorhanden ist. Dies bedeutet, dass solche Studien entweder in Kooperation etwa mit der empirischen Bildungsforschung oder unter Einbezug von Mitarbeiterinnen und Mitarbeitern etwa aus der Psychologie oder der Soziologie durchgeführt werden müssen.

Diese Einschränkung kann jedoch nicht als prinzipielles Hindernis angesehen werden. Wirksamkeitsuntersuchungen stehen aber zusätzlich vor einem *Dilemma*, das nicht leicht aufzulösen ist. Wirksamkeit ist jeweils nur durch Vergleiche abzuschätzen, was mit sich bringt, dass bestimmte Einflussgrößen konstant gehalten werden müssen, weil sonst eine Vergleichbarkeit der Effekte in religionsdidaktischer Hinsicht nicht gegeben wäre. Anders ausgedrückt muss sichergestellt werden, dass es tatsächlich die für die Untersuchung ausgewählte Unterrichtsstrategie ist, von der bestimmte Wirkungen ausgehen. Bei Interventionsstudien bedeutet das etwa, dass das Treatment oder eine entsprechende Unterrichtseinheit vorgegeben und in verschiedenen Unterrichtssituationen in gleicher Weise eingesetzt werden muss (*treatment fidelity*). Genau diese Forderung widerspricht jedoch dem pädagogischen Alltagsverstand, der gerade umgekehrt eine *flexible Anpassung an unterschiedliche Unterrichtssituationen* verlangt. Das angesprochene Dilemma lässt sich also so fassen, dass anzustrebende flexible Unterrichtsformen sich solchen Designs eher zu entziehen scheinen. Eigentlich soll der „normale“ Unterricht untersucht werden, faktisch wird aber ein nicht (ganz) alltäglicher Unterricht erfasst. Damit werden die Befunde aber nicht wertlos, besonders wenn es gelingt, beispielsweise bei der Entwicklung von Treatments auch erfahrene Praktikerinnen und Praktikern einzubeziehen. Auf diese Weise kann zumindest versucht werden, den Abstand zwischen dem Alltag von Unterricht und der Untersuchungssituation zu verringern. Bei den Tübinger Interventionsstudien wurde deshalb konsequent so verfahren, dass die Aufgabe der Treatmententwicklung insbesondere bei Kolleginnen und Kollegen aus der Praxis lag.

5 Wirksamkeitsuntersuchungen als Forschungsformat?

In einem letzten Schritt soll nun die mit dieser Ausgabe von Theo-Web verbundene Frage aufgenommen werden, ob sich die religionsdidaktische Wirksamkeitsforschung als ein „*Forschungsformat*“ verstehen lässt. Die Antwort auf diese Frage setzt allerdings zunächst eine Klärung des Begriffs der Forschungsformate voraus.

Der Begriff des Forschungsformats begegnet im vorliegenden Zusammenhang vor allem im Rahmen der neuen Diskussion über die Allgemeine Fachdidaktik, insbesondere in der Gesellschaft für Fachdidaktik (GFD, Bayrhuber et al., 2012; Rothgangel et al., 2020). Dort wird er nicht im Rückgriff auf eine bestimmte Begriffs- oder Forschungstradition verwendet, sondern als eine neue begriffliche Konstruktion, zu der noch kein Konsens vorauszusetzen ist (Rothgangel, 2020, S. 535). Als maßgeblich zitiert wird folgende Definition aus einem Positionspapier der GFD aus dem Jahr 2016: Ein „Format fachdidaktische Forschung“ bezeichnet „die Gesamtheit aller inhaltlichen, methodischen und forschungsorganisatorischen Aspekte [...], die bei der Planung, Durchführung, Auswertung und Ergebnisverwertung eines fachdidaktischen Forschungsvorhabens beschrieben werden können“ (Riegel & Rothgangel, 2020a, S. 350).

Bei dem hier beschriebenen Ansatz der religionsdidaktischen Wirksamkeitsforschung hat sich besonders das Verhältnis zwischen Inhalten und Methoden oder Designs als wesentlich erwiesen. Die *methodischen Zugriffsweisen*, die im Design eines Forschungsprojekts eingesetzt werden, müssen immer wieder auf die Inhalte des Religionsunterrichts abgestimmt werden. Umgekehrt müssen dabei die vor allem in der empirischen Bildungsforschung und der Pädagogischen Psychologie maßgeblichen Standards berücksichtigt werden, wobei weder bei der einen Seite – den Inhalten – noch bei der anderen Seite – den Standards der empirischen Forschung – Abstriche gemacht werden können. Von „forschungsorganisatorischen Aspekten“ lässt sich dabei wohl vor allem im Blick auf die erforderlichen Kooperationen sprechen, auf die insbesondere bei Interventionsstudien hinzuweisen war. Im Sinne der in der obigen Definition genannten „Ergebnisverwertung“ könnte auch noch an die Rezeption der Forschungsbefunde gedacht werden, bei der dafür gesorgt werden muss, dass die Praxis im Unterrichtsfach Religion tatsächlich erreicht werden kann. Dafür sollten eigene Kommunikationswege vorgesehen werden, ohne die ein praxisorientiertes Forschungsprojekt nicht an sein Ziel kommt.

Mit dem Zusammenspiel inhaltlicher, methodischer und forschungsorganisatorischer Aspekte bringt der Begriff der fachdidaktischen Forschungsformate so gesehen tatsächlich eine Betrachtungsweise ins Spiel, die für fachdidaktische Forschung spezifisch ist. Denn gleichsam selbstverständlich kreist dieser Forschungsbereich um das Verhältnis zwischen bestimmten inhaltlichen Domänen auf der einen und empirischen Zugangsweisen auf der anderen Seite.

Da der Begriff der Formate von Forschung der Diskussion der GFD entstammt, ist verständlich, warum der Begriff vor allem den besonderen Charakter *fachdidaktischer* Forschung sowie deren Eigenständigkeit im Vergleich zu anderen Bereichen der Forschung hervorheben soll. Nicht zuletzt sind dabei auch mögliche Drittmittelgeber im Blick, etwa die Deutsche Forschungsgemeinschaft (DFG). Die Eigenbedeutung fachdidaktischer Forschung hervorzuheben kann auch im Blick auf die religionsdidaktische Wirksamkeitsforschung als sinnvoll bezeichnet werden, eben im Blick auf das beschriebene Ineinandergreifen inhaltlicher bzw. domänenspezifischer Aspekte sowie Verfahrensweisen aus der empirischen Bildungsforschung bzw. der Pädagogischen Psychologie. Insofern stellt die religionsdidaktische Wirksamkeitsforschung in der Tat ein eigenes Format dar, eben im Sinne fachdidaktischer Forschung und im Unterschied zu Untersuchungen, die sich nicht auf das Lehren und Lernen in einem bestimmten Fach beziehen.

Gefragt werden muss allerdings, in welchem Sinne diese Besonderheit des fachlichen Bezugs zu verstehen ist. Denn die religionsdidaktische Wirksamkeitsforschung stellt eine direkte oder auch indirekte Form der Kooperation mit der empirischen Bildungsforschung dar. Dabei gelten, wie gesagt, die für die empirische Bildungsforschung maßgeblichen Kriterien auch für die religionsdidaktische Wirksamkeitsforschung. Auch wenn die Frage der Wirksamkeit auf fachbezogene Lern- oder Bildungsprozesse bezogen wird, kommt es ja entscheidend auf die für die empirische Bildungsforschung maßgeblichen Kriterien etwa der Reliabilität und der Validität an. Ohne verlässliche Vergleichsmöglichkeiten sind auch in der Fachdidaktik keine Aussagen zur Wirksamkeit möglich. Die Fachlichkeit gestaltet den Umgang mit Validitätsprüfungen inhaltlich aus, aber sie verändert diese als solche nicht.

Wie Riegel und Rothgangel in der Einleitung zu dieser Ausgabe von Theo-Web deutlich machen, soll der Formatbegriff auch bessere Voraussetzungen für eine *Beschreibung*

religionspädagogischer Forschung bieten als die in ihrer Sicht herkömmliche Einteilung nach Methoden. Der Diskussion über fachdidaktische Forschungsformate ist allerdings bereits innerhalb der GFD anzumerken, dass die Unterscheidung zwischen Formaten und Methoden fließend ist. In der Einleitung zu dem GFD-Band „Formate fachdidaktischer Forschung“ verweist Lutz-Helmut Schön (2012) als damaliger Vorsitzender der GFD mit seiner eigenen Formulierung auf ein Verständnis, das sich kaum vom Methodenbegriff unterscheidet: „Die skizzierten Pole fachdidaktischer Forschungsansätze bedingen sehr unterschiedliche Forschungsformate, die von der qualitativen Inhaltsanalyse nach Mayring über kategorienbasierte Videoanalysen realen Unterrichts bis zu strengen, klinischen Analysen im Sinne der empirischen Sozialforschung reichen.“ (S. 7) Demnach bleibt die Frage nach Methoden auch dann bedeutsam, wenn von Formaten her gedacht wird. Riegel und Rothgangel kann insofern zwar gefolgt werden, wenn sie fordern, unterschiedliche Ansätze der religionsdidaktischen Forschung nicht einfach von verschiedenen Methoden her zu charakterisieren. Das entspricht auch dem oben berichteten Befund, dass heute häufig mit Mixed-Methods-Ansätzen gearbeitet wird. Allerdings machen Riegel und Rothgangel nicht genügend deutlich, dass durch ihren Einwand die Frage nach Methoden der religionspädagogischen Forschung selbst keineswegs obsolet wird. Der Einwand trifft lediglich einen bestimmten Gebrauch des Methodenbegriffs, nicht aber den Sinn der Aufgabe, sich mit Methodenfragen auseinanderzusetzen. In meiner eigenen Darstellung zur religionspädagogischen bzw. -didaktischen Forschung habe ich mich zwar auf Methoden konzentriert, daraus aber keine Einteilung dieser Forschung abgeleitet (Schweitzer, 2019, S. 263–286). Religionsdidaktisch interessant erscheint auch die Möglichkeit, Untersuchungen dem Angebots-Nutzungs-Modell zuzuordnen (Schweitzer, 2020).

Wie weit reicht umgekehrt der Versuch, die religionsdidaktische Forschung mit Hilfe des Formatbegriffes zu beschreiben und zu typisieren? Können in der religionsdidaktischen Forschung distinkte Einzelformate identifiziert werden oder begegnen hier sich überlappende Ansätze? Am Beispiel der Unterscheidung zwischen religionsdidaktischer Wirksamkeitsforschung und religionsdidaktischer Entwicklungsforschung hat sich in der vorliegenden Darstellung gezeigt, dass beide kaum trennscharf auseinanderzuhalten sind. Die Entwicklungsforschung setzt in entscheidender Hinsicht eine Evaluation der entwickelten Unterrichtseinheiten voraus, ganz im Sinne der Wirksamkeitsforschung. Und zumindest in gewisser Weise konvergieren Wirksamkeits- und Entwicklungsforschung im Gedanken der Intervention.

Solche Beobachtungen könnten dafür sprechen, dass es für den Gebrauch des Formatbegriffs vor allem auf die Klärung des Eigencharakters fachdidaktischer Forschung ankommt, also gleichsam im *Außenverhältnis* zu anderen Formen der Forschung, und entsprechend *weniger* auf eine Ausdifferenzierung von Forschungsformaten *innerhalb der Religionsdidaktik* selbst. Denn vielfach lassen sich Klärungen durch Forschung erst erreichen, wenn unterschiedliche religionsdidaktische Forschungsansätze miteinander kombiniert werden, beispielsweise also religionsdidaktische Entwicklungsforschung und religionsdidaktische Wirksamkeitsforschung.

Bei alledem muss auch die grundlegende Einbindung empirisch-religionsdidaktischer Forschung in weitere Horizonte bewusst bleiben. Empirisch-religionsdidaktische Forschung steht nicht für sich allein, sondern bedarf analytisch-hermeneutischer sowie historischer und vergleichender Perspektiven, die es allererst erlauben, empirische Befunde angemessen zu interpretieren. In anderer Weise sind solche Perspektiven als Voraussetzung für empirische Forschung in der Religionsdidaktik anzusprechen, weil entsprechende Fragestellungen für empirische Untersuchungen aus ihnen hervorgehen.

Literaturverzeichnis

- Bayrhuber, H., Harms, U., Muszynski, B., Ralle, B., Rothgangel, M., Schön, L.-H., Vollmer, H.J. & Weigand, H.-G. (Hrsg.) (2012). *Formate fachdidaktischer Forschung. Empirische Projekte – historische Analysen – theoretische Grundlagen*. Münster & New York: Waxmann.
- Benner, D., Schieder, R., Schluß, H. & Willems, J. (Hrsg.) (2011). *Religiöse Kompetenz als Teil öffentlicher Bildung. Versuch einer empirisch, bildungstheoretisch und religionspädagogisch ausgewiesenen Konstruktion religiöser Dimensionen und Anspruchsniveaus*. Paderborn: Schöningh.
- Conroy, J. C., Lundie, D., Davis, R. A., Baumfield, V., Barnes, L. P., Gallagher, T., Lowden, K., Bourque, N. & Wenel, K. (2013). *Does Religious Education work? A Multi-dimensional Investigation*. New York: Bloomsbury Academic.
- Francis, L., Penny, G. & McKenna, U. (2016). Does RE work and contribute to the common good in England? In E. Arweck (Hrsg.), *Young People's Attitudes to Religious Diversity* (S. 153–169). London & New York: Routledge.
- Gärtner, C. (Hrsg.) (2018a). *Religionsdidaktische Entwicklungsforschung. Lehr-Lernprozesse im Religionsunterricht initiieren und erforschen*. Stuttgart: Kohlhammer.
- Gärtner, C. (2018b). Einführung in Fachdidaktische Entwicklungsforschung aus religionsdidaktischer Perspektive. In C. Gärtner (Hrsg.), *Religionsdidaktische Entwicklungsforschung. Lehr-Lernprozesse im Religionsunterricht initiieren und erforschen* (S. 11–30). Stuttgart: Kohlhammer.
- Gärtner, C. & Bettin, N. (Hrsg.) (2015). *Interreligiöses Lernen an außerschulischen Lernorten. Empirische Erkundungen zu didaktisch inszenierten Begegnungen mit dem Judentum*. Berlin: LIT.
- Göllner, R., Wagner, W., Klieme, E., Lüdtke, O., Nagengast, B. & Trautwein, U. (2016). Erfassung der Unterrichtsqualität mithilfe von Schülerurteilen: Chancen, Grenzen und Forschungsperspektiven. In BMBF (Hrsg.), *Forschung in Anknüpfung an Large-Scale Assessments*. Bonn: Bundesministerium für Bildung und Forschung (S. 63–82). (Bildungsforschung Band 44). Bielefeld: Bertelsmann Verlag.
- Göbinger, C. (2014). *Die Rezeption der Schöpfungserzählung nach Gen 1-2,4a bei Grundschulern. Eine fächerübergreifende Interventionsstudie*. (Dissertation) Frankfurt a. M.: Lang.
- Ilg, W., Pohlers, M., Gräbs Santiago, A. & Schweitzer, F. (2018). *Jung – Evangelisch – Engagiert. Langzeiteffekte der Konfirmandenarbeit und Übergänge in ehrenamtliches Engagement. Empirische Studien im biografischen Horizont*. Gütersloh: GVH.
- Ivanov, S. (2016). Eine Evaluationsstudie in NRW. In D. Benner & R. Nikolova (Hrsg.), *Ethisch-moralische Kompetenz als Teil öffentlicher Bildung* (S 187–194). Paderborn: Ferdinand Schöningh.

- Klieme, E. et al. (2007). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*, hg. v. Bundesministerium für Bildung und Forschung. Bonn: BMBF.
- Kohlberg, L. (1995). *Die Psychologie der Moralentwicklung*. Frankfurt am Main: Suhrkamp.
- Losert, M., Merkt, H. & Schweitzer, F. (2015). In Search of Interreligious Competence: An Empirical Study in the Context of Training Caregivers Through Religious Education. *Journal of Empirical Theology*, 28, S. 90–112.
- Meyer, K. (2019). *Grundlagen interreligiösen Lernens*. Göttingen: V&R.
- Obst, G. (2008). *Kompetenzorientiertes Lehren und Lernen im Religionsunterricht (4. Aufl. überarb. und aktualisiert von H. Lenhard 2015)*. Göttingen: V&R.
- Oser, F. (1988). *Wieviel Religion braucht der Mensch? Erziehung und Entwicklung zur religiösen Autonomie*. Gütersloh: GVH.
- Prediger, S. (2018). Religionsdidaktische Entwicklungsforschung. Kommentar zu Chancen und Grenzen eines Forschungsformats. In C. Gärtner (Hrsg.), *Religionsdidaktische Entwicklungsforschung. Lehr-Lernprozesse im Religionsunterricht initiieren und erforschen* (S. 165–174). Stuttgart: Kohlhammer.
- Rzejak, D. & Lipowsky, F. (2019). *Abschlussbericht zur wissenschaftlichen Begleitung der Fortbildung „Vielfalt fördern“*. Kassel: Universität Kassel.
- Riegel, U. & Kindermann, K. (2017). *Field Trips to the Church: Theoretical Framework, Empirical Findings, Didactic Perspectives*. Münster & New York: Waxmann.
- Riegel, U. & Rothgangel, M. (2020). Religionsdidaktik. In M. Rothgangel, U. Abraham, H. Bayrhuber, V. Frederking, W. Jank & H. J. Vollmer (Hrsg.), *Lernen im Fach und über das Fach hinaus. Bestandsaufnahmen und Forschungsperspektiven aus 17 Fachdidaktiken im Vergleich. Allgemeine Fachdidaktik Bd.2* (S. 339–362). Münster & New York: Waxmann. [= 2020a]
- Riegel, U. & Rothgangel, M. (2020). Das „Format fachdidaktischer Forschung“ als heuristischer Begriff und seine religionsdidaktische Bedeutung. *TheoWeb*, 19(1), 2–16.[= 2020b]
- Ritzer, G. (2010). *Interesse – Wissen – Toleranz – Sinn. Ausgewählte Kompetenzbereiche und deren Vermittlung im Religionsunterricht. Eine Längsschnittstudie*. Wien: LIT.
- Rodegro, M. (2010). *Schöpfung oder Urknall? Eine empirische Untersuchung im Religionsunterricht der Sekundarstufe II*. Kassel: Kassel University Press.
- Rothgangel, M. (2020). 17 Fachdidaktiken im Vergleich. Ergebnisse und Perspektiven. In M. Rothgangel, U. Abraham, H. Bayrhuber, V. Frederking, W. Jank & H. J. Vollmer (Hrsg.), *Lernen im Fach und über das Fach hinaus. Bestandsaufnahmen und Forschungsperspektiven aus 17 Fachdidaktiken im Vergleich. Allgemeine Fachdidaktik Bd.2* (S. 469–580). Münster & New York: Waxmann.
- Rothgangel, M., Abraham, U., Bayrhuber, H., Frederking, V., Jank, W. & Vollmer, H.J. (Hrsg.) (2020). *Lernen im Fach und über das Fach hinaus. Bestandsaufnahmen und Forschungsperspektiven aus 17 Fachdidaktiken im Vergleich. Allgemeine Fachdidaktik Bd.2*. Münster & New York: Waxmann.
- Schleiermacher, F. (1849). *Erziehungslehre*. Berlin: Reimer.

- Schön, L.-H. (2012). GFD-Fachtagung 2011. Formate Fachdidaktischer Forschung. In H. Bayrhuber, U. Harms, B. Muszynski, B. Ralle, M. Rothgangel, L.-H. Schön, H. J. Vollmer & H.-G. Weigand (Hrsg.). *Formate fachdidaktischer Forschung. Empirische Projekte – historische Analysen – theoretische Grundlegungen* (S. 7–9). Münster & New York: Waxmann.
- Schulz, K.F., Altman, D.G. & Moher, D. for the CONSORT Group (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2844940/> [Zugriff: 26.04.2020].
- Schwarz, S. (2019). *SchülerInnenperspektiven und Religionsunterricht. Empirische Einblicke – Theoretische Überlegungen*. Stuttgart: Kohlhammer.
- Schweitzer, F. (2018). Interventionsstudien im Religionsunterricht: Begründung – Beispiele – Perspektiven. In M. Schambeck & U. Riegel (Hrsg.), *Was im Religionsunterricht so läuft. Wege und Ergebnisse religionspädagogischer Unterrichtsfor- schung* (S. 107–122). Freiburg: Herder.
- Schweitzer, F. (2019). *Religionspädagogik*. Gütersloh: GVH.
- Schweitzer, F. (2020). *Religion noch besser unterrichten. Qualität und Qualitätsentwicklung im RU*. Göttingen: V&R.
- Schweitzer, F. & Boschki, R. (Hrsg.) (2018). *Researching Religious Education. Classroom Processes and Outcomes*. Münster & New York: Waxmann.
- Schweitzer, F., Bräuer, M. & Boschki, R. (Hrsg.) (2017). *Interreligiöses Lernen durch Perspektivenübernahme. Eine empirische Untersuchung religionsdidaktischer Ansätze*. Münster & New York: Waxmann.
- Schweitzer, F. & Bucher, I. (Hrsg.) (2020). *Judentum und Islam im Religionsunterricht. Theoretische Analysen und empirische Befunde im Horizont interreligiösen Lernens*. Münster & New York: Waxmann.
- Schweitzer, F., Ilg, W. & Schreiner, P. (Hrsg.) (2019). *Researching Non-Formal Religious Education in Europe*. Münster & New York: Waxmann.
- Simojoki, H., Ilg, W., Schlag, T. & Schweitzer, F. (2018). *Zukunftsfähige Konfir- mandenarbeit. Empirische Erträge – theologische Orientierungen – Perspektiven für die Praxis*. Gütersloh: GVH.
- Wagensommer, G. & Schweitzer, F. (Hrsg.) (2018). *Wertebildung, Interesse und Reli- gionsunterricht. Ethisch und religiös ausgerichteter Unterricht im Vergleich. The- oretische und empirische Untersuchungen zur Wirksamkeit des BRU*. Münster & New York: Waxmann.
- Ziebertz, H.-G. (Hrsg.) (2010). *Gender in Islam und Christentum. Theoretische und empirische Studien*. Berlin: LIT.

Dr. Friedrich Schweitzer, Professor für Praktische Theologie/Religionspädagogik an der Evangelisch-Theologischen Fakultät der Universität Tübingen.