

Gärtner, Holger

Herausforderungen und Gelingensbedingungen für die Umsetzung (quasi-)experimenteller Untersuchungsdesigns im Forschungsfeld Schule

Die Deutsche Schule 118 (2026) 2, S. 108-119



Quellenangabe/ Reference:

Gärtner, Holger: Herausforderungen und Gelingensbedingungen für die Umsetzung
(quasi-)experimenteller Untersuchungsdesigns im Forschungsfeld Schule - In: Die Deutsche Schule 118
(2026) 2, S. 108-119 - URN: urn:nbn:de:0111-pedocs-357912 - DOI: 10.25656/01:35791;
10.31244/dds.2026.02.04

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-357912>

<https://doi.org/10.25656/01:35791>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

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

Mit der Verwendung dieses Dokuments erkennen Sie die
Nutzungsbedingungen an.

Terms of use

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

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



Kontakt / Contact:

peDOCS

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Informationszentrum (IZ) Bildung

E-Mail: pedocs@dipf.de

Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Holger Gärtner

Herausforderungen und Gelingensbedingungen für die Umsetzung (quasi-)experimenteller Untersuchungsdesigns im Forschungsfeld Schule

Zusammenfassung

Die empirische Bildungsforschung in Deutschland ist stark gewachsen, dennoch besteht weiter Bedarf an Erklärungswissen zu wirksamen Förder- und Implementationsansätzen. (Quasi-)Experimentelle Interventionsstudien ermöglichen Aussagen über Wirksamkeit und Effizienz von Maßnahmen, genau diese sind im Schulbereich aber selten. Der Beitrag diskutiert Herausforderungen und Gelingensbedingungen bei der Durchführung solcher Studien sowie Impulse durch Forschungsförderung und Kooperationen mit Landesinstituten.

Schlüsselwörter: Forschungsdesign; Quasi-Experiment; Forschungsförderung; Kooperation; Maßnahme

Barriers and Enabling Conditions for Implementing (Quasi-)Experimental Designs in School Research

Abstract

Empirical educational research in Germany has grown significantly, yet there is still a need for explanatory knowledge about effective measures and implementation strategies. (Quasi-)Experimental intervention studies enable conclusions to be drawn about the effectiveness and efficiency of measures, but these are rare in the school sector. This article discusses barriers and enabling conditions for conducting such studies, as well as impulses provided by research funding and cooperation with state institutes.

Keywords: research design; quasi-experiment; research funding; co-operation; measure

1 Einleitung

Die im Jahr 2001 veröffentlichten PISA-Ergebnisse machten deutlich, dass die Leistungen der Schüler*innen in Deutschland im internationalen Vergleich unterdurchschnittlich waren und in erheblichem Ausmaß mit der sozialen Herkunft zusammenhingen, und sie zeigten

zugleich, dass bis dahin nur in begrenztem Umfang systematisches Wissen über die Leistungsfähigkeit dieses für die Gesellschaft zentralen Systems vorlag (Baumert et al., 2001). Seitdem kann ein rasanter Aufschwung der empirischen Bildungsforschung in Deutschland wahrgenommen werden (Zapp & Powell, 2017). So stehen in Deutschland ein Vierteljahrhundert später Forschungsstrukturen zur Verfügung, die regelmäßig über die Leistungsfähigkeit des Bildungssystems informieren (Gärtner, 2024). Grundlage hierfür ist nach wie vor die Strategie zum Bildungsmonitoring der Kultusministerkonferenz (KMK, 2006).

Bei all diesen Vorhaben handelt es sich um Beobachtungsstudien, die in erster Linie *beschreibende* Aussagen über Leistungsstände und deren Entwicklung ermöglichen. Entsprechend hat die KMK in ihrer Überarbeitung der Strategie zum Bildungsmonitoring 2015 betont, dass neben dem vorhandenen Beschreibungswissen vertieftes Erklärungswissen generiert werden soll, also Wissen darüber, warum es zu bestimmten Entwicklungen kommt und welche Maßnahmen geeignet sind, diagnostizierte Schwächen zu adressieren (KMK, 2015).

Thiel (2014) beschreibt unterschiedliche Typen wissenschaftlicher Evidenz und ihrer Nutzung in Entscheidungsprozessen. Für ein *Agenda-Setting* werden z. B. detaillierte Beschreibungen von Problemlagen, für die Phase der *Zielformulierung* und der konkreten *Entscheidung* über das weitere Vorgehen Wissen über potenzielle Maßnahmen (bestenfalls inkl. Wissen über Kosten-Nutzen-Abwägungen und potenzielle Risiken) benötigt. Während der Phase der *Umsetzung* ist Wissen über bestehende Implementationsbarrieren hilfreich, während in einer abschließenden *Evaluationsphase* Wissen über die Zielerreichung notwendig ist, woraus gegebenenfalls weitere Optimierungen der Maßnahme abgeleitet werden können. Dieser Beitrag beschäftigt sich mit der Evidenzgrundlage für die Phase der Entscheidungen über das weitere Vorgehen, in der Wissen über die Ausgestaltung und Wirksamkeit verschiedener Maßnahmen zum möglichen Umgang mit festgestellten Problemlagen, deren potenzielle Nebenwirkungen und deren Kosten-Nutzen-Verhältnis von besonderer Relevanz sind. Interventionsstudien, die (quasi-)experimentelle Designs einsetzen, stellen eine Möglichkeit dar, belastbare Aussagen über Wirkungen von Maßnahmen generieren zu können (Cook et al., 2002; Strohmmer & Wirtz, 2022). Gillon et al. (2024) zeigen beispielhaft, wie quasi-experimentelle Wirksamkeitsstudien innerhalb einer mehrphasigen Interventionskonzeption zu einer raschen Ausweitung eines neuseeländischen Leseförderungsprogramms führten.

2 Geringer Anteil an Interventionsstudien

Im Folgenden wird unter pädagogischer Intervention jede gezielte Veränderung in einem System verstanden (Leutner, 2013). Pädagogische Interventionsstudien liegen dann vor, wenn es „um die Prüfung gezielter pädagogischer Maßnahmen auf deren erwartete pädagogische Wirkung“ geht (ebd., S. 19). Laut Döring (2023) handelt es sich hierbei um explanative Studien mit hypothesenprüfendem Charakter (ebd., S. 195). Experimentelle Designs ermöglichen durch zufällige Gruppenzuteilung kausale Aussagen, quasi-experimentelle Designs arbeiten ohne eine zufällige Gruppeneinteilung (was die kausale Interpretation einschränkt). Nicht-experimentelle Studien werden im Folgenden nicht betrachtet.

Obwohl (quasi-)experimentelle Designs seit Langem verfügbar sind und in anderen Disziplinen – etwa in Teilen der Medizin oder der Psychologie – eine zentrale Rolle spielen,

verbleibt ihr Anteil in der empirischen Bildungsforschung insgesamt begrenzt. Weishaupt und Rittberger (2012) schätzen in ihrer Übersicht zur Bildungsforschung in Deutschland, dass der Anteil von Experimenten an allen im Zeitraum von 1998 bis 2007 durchgeführten Forschungsprojekten bei knapp 4 Prozent lag (ebd., S. 95). Während für Deutschland keine aktuelleren Daten bekannt sind, berichten international jüngst Brady et al. (2023) über Analysen aus Fachzeitschriften, die auf ähnliche Entwicklungen hindeuten. So wird für die pädagogische Psychologie über eine Abnahme des Anteils von Interventions- und experimentellen Studien in mehreren Fachzeitschriften informiert; der Anteil experimenteller Studien sank im betrachteten Zeitraum von 33 Prozent im Jahr 2000 auf 20 Prozent im Jahr 2020, der Anteil von Interventionsstudien von 40 Prozent auf 25 Prozent. Gopalan et al. (2020) zeigen für den Zeitraum von 1995 bis 2018 zwar einen relativen Anstieg von Veröffentlichungen mit quasi-experimentellem Design, der Anteil bleibt aber auch dort begrenzt.

Diese Befunde werden unterschiedlich bewertet: Während einerseits auf die Bedeutung kausaler Evidenz für Entscheidungen verwiesen wird (Leutner, 2013; Maag Merki, 2021; Strohmmer & Wirtz, 2022), wird in anderen Debatten auf die Grenzen experimenteller Designs in komplexen pädagogischen Settings und auf die Notwendigkeit verwiesen, ein plural angelegtes Methodenrepertoire zu erhalten (Cohen et al., 2018, Kap. 6; Connolly et al., 2018). In diesem Beitrag wird der Fokus auf (quasi-)experimentelle Interventionsstudien gelegt, da diese Studien dazu dienen, die Wirksamkeit sowie die Effizienz von Maßnahmen zu überprüfen, und sie so einen spezifischen Beitrag innerhalb einer Interventionskonzeption leisten. Darüber hinaus stellen einzelne Interventionsstudien die Grundlage zur systematischen Bündelung von Erkenntnissen über die Wirksamkeit von Maßnahmen dar (Leutner, 2013). Entsprechende Meta-Analysen und Überblicksarbeiten werden auch im deutschen Sprachraum zunehmend erstellt (Hetmanek et al., 2023) und schaffen die Möglichkeit, den Einfluss relevanter Rahmenbedingungen auf die Wirksamkeit einer Maßnahme zu untersuchen.

Nachfolgend werden Herausforderungen bei der Realisierung von Interventionsstudien mit (quasi-)experimentellen Designs skizziert, bevor anschließend potenzielle Gelingensbedingungen, insbesondere mit Blick auf Kooperationen und Forschungsförderung, vertiefend dargestellt werden.

3 Herausforderungen in der Realisierung von Interventionsstudien mit (quasi-)experimentellen Designs

Mittlerweile gibt es zahlreiche erfolgreiche Umsetzungen experimenteller und quasi-experimenteller Designs zur Untersuchung pädagogischer Interventionen auf unterschiedlichen Ebenen. Beispielhaft genannt seien:

- auf Unterrichtsebene die experimentelle Untersuchung der Wirksamkeit von Zeitzeugen im Geschichtsunterricht (Bertram et al., 2017);
- auf Schulebene die Untersuchung der Wirksamkeit von prosozialen Maßnahmen (Pas et al., 2019);
- auf Ebene von Verwaltungseinheiten die Untersuchung der Wirksamkeit evidenzbasierter Entscheidungen durch Schulaufsicht (Slavin et al., 2013).

Dennoch ist der Anteil solcher Studien im Feld Schule weiterhin begrenzt. Im Folgenden werden ausgewählte Gründe für diese Situation auf drei Ebenen diskutiert: struktureller Aufwand, Risiken im Projektverlauf und Überzeugungen zentraler Gruppen von Akteur*innen.

3.1 Aufwand

Vor dem Hintergrund hoher Belastungen im schulischen Alltag und einer angespannten Personalsituation fällt es Forschungsprojekten häufig schwer, genügend Lehrkräfte und Schulen für eine Teilnahme zu gewinnen, insbesondere wenn diese, wie oft bei Interventionsstudien, mit zusätzlichem Zeitaufwand verbunden ist (von Maurice et al., 2025). Hinzu kommen organisatorische Anforderungen zur Abstimmung mit und Einbindung von Schulleitungen, Schulaufsicht, schulischen Mitwirkungsgremien oder Erziehungsberechtigten.

Entsprechend wird der Aufwand häufig als ein Grund für den geringen Anteil an Interventionsstudien mit (quasi-)experimentellen Designs genannt (Brady et al., 2023; Maag Merki, 2021). Interventionsstudien sind dabei nicht per se aufwändiger als Beobachtungsstudien, beide Vorgehensweisen teilen viele Prozessschritte, etwa die Planung und Rekrutierung der Stichprobe, die Entwicklung bzw. Auswahl von Erhebungsinstrumenten, die behördliche Genehmigung, Datenschutzfragen sowie Datenerhebung und -auswertung (Döring, 2023, S. 24 f.). Der zentrale Unterschied liegt darin, dass Beobachtungsstudien den bestehenden Gang der Dinge abbilden, d. h. es werden natürliche Gruppen untersucht, ohne dass systematisch in das Handeln der beteiligten Akteur*innen eingegriffen wird. In Interventionsstudien ist demgegenüber typischerweise vorgesehen, dass Akteur*innen – zumeist Lehrkräfte, teilweise aber auch weitere Professionen – eine zusätzliche oder veränderte Maßnahme in ihrem Alltag umsetzen, etwa eine Fortbildung besuchen und darauf aufbauend ihren Unterricht über einen festgelegten Zeitraum anpassen.

Für die Begleitforschung zu einer Intervention ist eine geeignete Vergleichs- oder Kontrollgruppe erforderlich, um Effekte der Maßnahme möglichst zuverlässig von anderen Einflüssen unterscheiden zu können (Cook et al., 2002, S. 6). Dies kann etwa durch Randomisierung, durch vergleichbare Kontrollgruppen mit statistischer Kontrolle oder durch andere Formen quasi-experimenteller Kontrastierung erfolgen; in allen Fällen steigt der Koordinationsaufwand im Feld.

Insgesamt lässt sich festhalten, dass Interventionsstudien mit (quasi-)experimentellen Designs in zentralen Prozessschritten besondere Anforderungen stellen, insbesondere was Rekrutierung, Abstimmung und kontinuierliche Begleitung der teilnehmenden Schulen betrifft, und dass diese Anforderungen in angespannter Ressourcenlage deutlich spürbar werden (Brady et al., 2023; Demby et al., 2020).

3.2 Risiken

Interventionsstudien können prinzipiell in unterschiedlich großem Umfang durchgeführt werden; kleinere Vorhaben lassen sich teilweise über vorhandene Ressourcen von Lehrstühlen, Arbeitsbereichen oder Schulen realisieren, etwa im Kontext von Qualifikationsarbeiten. Häufiger werden größere Interventionsstudien jedoch im Rahmen projektförmeriger Forschungsförderung umgesetzt, z. B. durch die Deutsche Forschungsgemeinschaft (DFG), durch Ministerien oder durch Stiftungen (Zapp & Powell, 2017; Maag Merki, 2021).

Projektförderung ist in der Regel zeitlich befristet und an einen klaren Arbeits- und Maßnahmenplan gebunden; nach Projektende lösen sich viele Projektkonstellationen auf. Daraus ergeben sich mehrere Risiken für Interventionsstudien: Es kann etwa nicht gelingen, in der vorgesehenen Zeit ausreichend viele Schulen, Lehrkräfte oder Schüler*innen zu gewinnen, oder die Teilnahme fällt geringer aus als geplant, sodass die angestrebten Fallzahlen für eine belastbare Wirkungsschätzung nicht erreicht werden (Cook et al., 2002; Demby et al., 2020).

Die Rekrutierung und langfristige Bindung von Kontroll- bzw. Vergleichsgruppen kann eine zusätzliche Herausforderung darstellen: Während sich ein Teil der an der Intervention beteiligten Akteur*innen aufgrund ihres inhaltlichen Interesses eher auch für die Teilnahme an der Begleitforschung gewinnen lässt, bedarf es bei Vergleichsgruppen häufig eines eigenen Kommunikations- und Vermittlungsaufwands, um eine kontinuierliche Beteiligung sicherzustellen. Ausfälle im Panel, insbesondere in Kontrollgruppen, sowie kleinere Stichproben erschweren die Generalisierbarkeit und Präzision der Wirksamkeitsschätzungen (Brady et al., 2023; Demby et al., 2020). Interventionsstudien auf freiwilliger Basis erfordern daher eine sorgfältige Beschreibung der Stichproben und Kontexte sowie Schlussfolgerungen darüber, für welche Zielgruppen und Rahmenbedingungen die untersuchten Maßnahmen als wirksam angesehen werden können (Cook et al., 2002).

Ein weiteres Risiko besteht darin, dass die Ergebnisse einer Interventionsstudie nicht in der gewünschten Form veröffentlicht werden können, z. B. aufgrund begrenzter Stichprobengrößen oder hoher Ausfälle. Vor dem Hintergrund, dass in vielen Wissenschaftssystemen die hochrangige Publikation von Projektergebnissen ein zentrales Kriterium für die Bewertung wissenschaftlicher Leistungen ist (Gärtner et al., 2023), kann dies dazu beitragen, dass Forschende die mit Interventionsstudien verbundenen Risiken gegenüber weniger risikobehafteten Studiendesigns abwägen.

Nicht zuletzt hängt die Teilnahmebereitschaft von Schulen und Lehrkräften auch von den Anreizen ab, die mit der Teilnahme verbunden sind. In vielen Fällen können Forschungsprojekte z. B. die Anerkennung von Teilen der Intervention als Fortbildung oder eine engere inhaltliche Begleitung bieten; darüberhinausgehende Entlastungen, etwa Deputatsreduzierungen, liegen hingegen in der Verantwortung von Schulleitungen und Schulverwaltung und sind in Zeiten knapper Personalressourcen schwer umzusetzen (von Maurice et al., 2025).

3.3 Überzeugungen

Die Rahmenbedingungen des schulischen Alltags und die Belastung des pädagogischen Personals prägen die Bereitschaft, an Interventionsstudien teilzunehmen, in erheblichem Maße (von Maurice et al., 2025). Darüber hinaus spielen Überzeugungen und Einschätzungen zur Relevanz von Bildungsforschung und zu bestimmten Forschungsansätzen eine wichtige Rolle.

Hartmann und Kunter (2022) berichten zusammenfassend, dass Lehrkräfte im Berufsalltag wissenschaftliche Befunde nur gelegentlich wahrnehmen und sich stärker auf eigene Erfahrungen oder kollegiale Empfehlungen stützen. Vor diesem Hintergrund ist nachvollziehbar, dass vielen Lehrkräften die Teilnahme an Forschungsstudien im Allgemeinen und an Interventionsstudien im Besonderen nicht unmittelbar als Teil ihrer beruflichen Rolle erscheint und dass Forschungsanfragen zunächst mit anderen Anforderungen des Alltags koordiniert werden müssen.

Auch in der wissenschaftlichen Community werden (quasi-)experimentelle Designs kontrovers diskutiert. Connolly et al. (2018) zeigen in ihrer systematischen Übersicht experimenteller Studien im Bildungsbereich sowohl Potenziale als auch Grenzen auf, Cohen et al. (2018) verweisen auf die Kontextabhängigkeit von Bildungsprozessen und problematisieren Erwartungen, komplexe Wirkmechanismen mit einzelnen Designs vollständig abbilden zu können. In der Folge gibt es in der Bildungsforschung neben Befürworter*innen experimenteller Ansätze auch Stimmen, die deren Aussagekraft für bestimmte Fragestellungen skeptisch beurteilen oder andere Zugänge priorisieren.

Zusammenfassend lässt sich damit festhalten, dass (quasi-)experimentelle Interventionsstudien im Forschungsfeld Schule nicht nur vor praktischen und strukturellen Herausforderungen stehen, sondern auch in einem Feld unterschiedlicher Überzeugungen und Erwartungen verortet sind – sowohl auf Seiten der Praxisakteur*innen als auch auf Seiten der Forschung.

4 Gelingensbedingungen

In diesem Beitrag wird davon ausgegangen, dass (quasi-)experimentelle Interventionsstudien einen spezifischen Beitrag zur empirischen Schulforschung leisten, indem sie Aussagen über die Wirksamkeit von Maßnahmen, über deren Effizienz und über Nebenwirkungen ermöglichen (Maag Merki, 2021; Strohmmer & Wirtz, 2022). Diese Form der Evidenz bildet einen Baustein in einem breiteren Spektrum von Wissensformen, der für Entscheidungen zur Weiterentwicklung von Unterricht und Schule herangezogen werden kann (Thiel, 2014), wie z. B. dergestalt, dass Schulen erfolgreiche Fortbildungs- bzw. Entwicklungsprogramme angeboten werden können (Gillon et al., 2024).

Von Maurice et al. (2025) benennen mehrere projektinterne Maßnahmen, die eine erfolgreiche Durchführung empirischer Studien an Schulen unterstützen können. Zentrale Bedeutung kommt dabei einer klaren Kommunikation zu, in der sowohl die Sinnhaftigkeit des Projekts als auch der pädagogische Mehrwert für die beteiligten Schulen herausgestellt werden. Von Beginn an sind zudem der zeitliche Aufwand für die Teilnahme

an der Intervention sowie an der Begleitforschung transparent zu machen. Weiterhin ist gegenüber den Schulen zu präzisieren, mit welchen Formen der Ergebnismeldung sie rechnen können und welche Unterstützung sie bei der Interpretation der Rückmeldungen erhalten. Darüber hinaus ist sicherzustellen, dass alle Datenerhebungen ökonomisch geplant und in Einklang mit den organisatorischen Rahmenbedingungen des schulischen Alltags umgesetzt werden.

Abschließend werden zur möglichen Steigerung des Anteils (quasi-)experimenteller Interventionsstudien zwei Ansatzpunkte diskutiert, die über eine projektinterne Optimierung der Kommunikation bzw. Unterstützung der Schulen hinausreichen: erstens die Frage, wie Verantwortung für die Durchführung von Interventionsstudien zwischen unterschiedlichen Akteur*innen aufgeteilt und kooperativ gestaltet werden kann, und zweitens die Rolle von Forschungsförderung bei der Unterstützung solcher Vorhaben.

4.1 Gemeinsame Verantwortung von Beginn an

Viele der beschriebenen Herausforderungen bei der Durchführung von Interventionsstudien im Feld Schule hängen damit zusammen, dass Forschungsprojekte bislang häufig weitgehend eigenverantwortlich versuchen, Zugang zu Schulen zu gewinnen, Rekrutierung und Kommunikation zu organisieren und gleichzeitig wissenschaftlichen Ansprüchen zu genügen. Demby et al. (2020) verdeutlichen am Beispiel von Evaluationen mit randomisierten Designs, wie bedeutsam partnerschaftliche Strukturen im Feld für die erfolgreiche Umsetzung sind.

In der deutschsprachigen Diskussion werden Landesinstitute und vergleichbare Einrichtungen seit einiger Zeit als wichtige Schnittstellen zwischen Wissenschaft, Politik und Praxis beschrieben (Bieber et al., 2018; Düllberg et al., 2023; Holtappels, 2019). Sie verantworten u. a. Fort- und Weiterbildung, entwickeln Materialien und beraten Schulen und Bildungsverwaltungen; daraus ergeben sich Chancen für eine engere Verzahnung mit Forschungsprojekten.

Für eine vertiefte Kooperation zwischen Wissenschaft und Landesinstituten bei der Planung und Umsetzung von Interventionsstudien lassen sich folgende Vorschläge formulieren. *Erstens* erleichtert sie den Zugang zum Feld, da Landesinstitute über etablierte Kommunikationskanäle zu Schulen, Schulleitungen und Lehrkräften verfügen. *Zweitens* unterstützt sie Rekrutierung und Anreizsetzung, weil Teile der Intervention als Fortbildung anerkannt oder in bestehende Qualifizierungsangebote eingebunden werden können. *Drittens* können in gemeinsamen Entwicklungsprozessen Forschungsziele und schulische Rahmenbedingungen besser aufeinander bezogen und Umsetzungshürden frühzeitig identifiziert werden. *Viertens* können, wenn Personal in Landesinstituten frühzeitig in Interventionskonzepte eingearbeitet wird, erfolgreiche Ansätze nach Projektende besser gesichert, adaptiert oder skaliert werden.

Große Forschungsverbünde wie „SchuMaS“ (vgl. Maaz & Marx, 2024), „QuaMath“ (Prediger et al., 2024) oder „SINUS-Transfer“ (Prenzel, 2019) zeigen in unterschiedlichen Varianten, dass eine langfristige, kooperative Struktur zwischen Wissenschaft, Landesinstituten und weiteren Praxispartner*innen die Chancen erhöhen kann, dass aus Projekten

heraus nachhaltige Impulse für Unterrichts- und Schulentwicklung entstehen. Insbesondere die Qualifizierung von Multiplikator*innen spielt hierbei eine große Rolle (Prediger et al., 2024).

Gleichzeitig ist mit Blick auf die Rolle der Landesinstitute zu betonen, dass sie keine neutralen Akteure sind, sondern in bildungspolitische und administrative Steuerungslogiken eingebunden bleiben (Bieber et al., 2018; Zapp & Powell, 2017). Eine stärkere Einbindung in Interventionsforschung sollte daher so gestaltet werden, dass wissenschaftliche Unabhängigkeit gewahrt bleibt und Entscheidungsbefugnisse, Verantwortlichkeiten und Erwartungen in transparenten Vereinbarungen geklärt werden (vgl. z. B. DFG-Vordruck 41.026 a-05/22¹). In einem solchen Verständnis wird die Verantwortung für Interventionsstudien nicht einseitig auf „die Wissenschaft“ verlagert, sondern zwischen unterschiedlichen Institutionen geteilt; gleichzeitig wird Wissenstransfer nicht als eindimensionale Bewegung von der Wissenschaft in die Praxis, sondern als wechselseitiger Austausch verstanden, in dem Erfahrungen aus der Praxis wiederum Forschungsfragen und -designs beeinflussen (Farley-Ripple et al., 2018). Gleichwohl gilt, dass sich die Anforderungen an Landesinstitute hierdurch verändern würden und sich dies in entsprechenden Organisationsplänen und Ressourcenzuweisungen niederschlagen müsste (Bieber et al., 2018).

4.2 Veränderte Forschungsförderung

Ein weiterer Ansatzpunkt zur Stärkung von Interventionsstudien und ihrer Nachhaltigkeit liegt in der Ausgestaltung von Forschungsförderung. Zapp und Powell (2017) dokumentieren, dass nach PISA erhebliche Mittel in die empirische Bildungsforschung in Deutschland geflossen sind. Die Drittmiteinnahmen im Rahmen der Erziehungswissenschaft haben sich von ca. 28 Millionen Euro im Jahr 2000 auf 110 Millionen Euro im Jahr 2021 erhöht (Schmidt-Hertha et al., 2024). Für die USA zeigen Gopalan et al. (2020), dass gezielte Förderlinien des Institute of Education Sciences (IES), in denen unter anderem (quasi-) experimentelle Evaluationsdesigns explizit adressiert werden, zu einem Anstieg entsprechender Studien beigetragen haben. Hedges (2018) hebt die Bedeutung begleitender Qualifizierungsmaßnahmen hervor, weist aber zugleich darauf hin, dass Programme wie das zum IES gehörende „What Works Clearinghouse“ bestimmte Verständnisse von „usable knowledge“ privilegieren und damit auch Debatten über die Reichweite solcher Modelle auslösen. Zapp und Powell (2017) diskutieren ebenso potenzielle Spannungen zwischen forschungspolitischen Steuerungsansprüchen und wissenschaftlicher Autonomie.

Vor diesem Hintergrund lassen sich für die Ausgestaltung von Forschungsförderung folgende Vorschläge formulieren:

- 1) Zunächst sollten explizite Förderlinien für Interventionsstudien eingerichtet werden.
- 2) Zudem erscheint es sinnvoll, in Programmen zur Entwicklung und Implementierung neuer Maßnahmen mindestens einen quasi-experimentellen Evaluationsstrang zu verankern, sofern dies mit der Fragestellung vereinbar ist.
- 3) Ergänzend sollte die Sekundärnutzung vorhandener Daten, etwa aus Bildungsmonitoring oder Lernstandsdiagnosen, gefördert werden, um eigene Evaluationen zu ergänzen.

1 <https://www.dfg.de/de/formulare-41-026a-246594>

zen, wobei die Limitationen dieser Daten transparent zu berücksichtigen sind (von Maurice et al., 2025).

- 4) Schließlich ist es wichtig, Ressourcen für die frühzeitige Zusammenarbeit mit Schnittstelleninstitutionen in Konzeption, Rekrutierung, Umsetzung und Transfer bereitzustellen, einschließlich der Finanzierung von Stellen an der Schnittstelle zwischen Forschung und Praxis (Düllberg et al., 2023; Holzer et al., 2024).

Solche Förderkriterien können dazu beitragen, den Anteil an Interventionsstudien zu erhöhen und die Wahrscheinlichkeit zu verbessern, dass erfolgreiche Maßnahmen nach Projektende nicht isoliert bleiben, sondern in bestehende Strukturen der Aus- und Fortbildung sowie der Schulentwicklung eingebettet werden (Holtappels, 2019).

5 Ausblick

Rund 25 Jahre nach der Veröffentlichung der ersten PISA-Ergebnisse steht deutlich mehr Beschreibungswissen über das Bildungssystem zur Verfügung als zu Beginn der 2000er Jahre (Gärtner, 2024). Die Bedarfe an Erklärungs- und Handlungswissen zur Weiterentwicklung von Unterricht und Schule sind nach wie vor groß, da sich an den von PISA 2000 festgestellten Problemlagen wenig geändert hat (Lewalter et al., 2023).

In diesem Beitrag wurde argumentiert, dass (quasi-)experimentelle Interventionsstudien einen wichtigen Baustein in einem Spektrum empirischer Schulforschung darstellen, weil sie – bei allen Grenzen – kausale Aussagen über die Wirksamkeit und die Effizienz von Maßnahmen ermöglichen und damit eine spezifische Form von Orientierung für Entscheidungen bereitstellen (Maag Merki, 2021; Strohmer & Wirtz, 2022; Thiel, 2014). Es wurde skizziert, welchen praktischen, strukturellen und forschungsbezogenen Herausforderungen Projekte gegenüberstehen, die solche Designs im Feld Schule umsetzen wollen, und es wurden mögliche Gelingensbedingungen diskutiert. Im Mittelpunkt standen dabei zwei Ansatzpunkte: eine frühzeitige, reflektierte Kooperation zwischen Forschung, Landesinstituten und weiteren Praxispartner*innen sowie eine Forschungsförderung, die Interventionsstudien gezielt ermöglicht, ohne andere Formen der Wissensgenerierung zu schmälern (Bieber et al., 2018; Holzer et al., 2024; Zapp & Powell, 2017).

Wenn es gelingt, in der empirischen Schulforschung immer häufiger nachvollziehbare Beispiele dafür zu entwickeln, wie wissenschaftliche Erkenntnisse auch aus (quasi-)experimentellen Interventionsstudien zur nachhaltigen Verbesserung von Unterricht und Schulentwicklung beitragen (Gillon et al., 2024), kann dies die Wahrnehmung der Relevanz empirischer Forschung in Praxis und Öffentlichkeit weiter stärken. Damit können (quasi-)experimentelle Interventionsstudien einen zentralen Beitrag dazu leisten, die nach wie vor bestehenden, u. a. durch PISA sichtbaren Leistungsrückstände und Ungleichheiten nicht nur zu beschreiben, sondern durch überprüfbare Maßnahmen systematisch zu verringern.

Literatur und Internetquellen

- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., & Weiss, M. (Hrsg.). (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Leske & Budrich.
<https://doi.org/10.1007/978-3-322-83412-6>
- Bertram, C., Wagner, W., & Trautwein, U. (2017). Learning historical thinking with oral history interviews: A cluster randomized controlled intervention study of oral history interviews in history lessons. *American Educational Research Journal*, 54 (3), 444–484.
<https://doi.org/10.3102/0002831217694833>
- Bieber, G., Egyptian, E. L., Klein, G., Oechslein, K., & Pikowsky, B. (2018). *Positionspapier der Landesinstitute und Qualitätseinrichtungen der Länder zum Transfer von Forschungswissen*.
https://ibbw-bw.de/site/pbs-bw-km-root/get/documents_E1235815622/KULTUS.Dachmandant/KULTUS/Dienststellen/ibbw/Empirische%20Bildungsforschung/Schwerpunkte/Wissenschaftstransfer/Positionspapier_Transfer_31.10.18.pdf
- Brady, A. C., Griffin, M. M., Lewis, A. R., Fong, C. J., & Robinson, D. H. (2023). How Scientific Is Educational Psychology Research? The Increasing Trend of Squeezing Causality and Recommendations from Non-intervention Studies. *Educational Psychology Review*, 35 (1).
<https://doi.org/10.1007/s10648-023-09759-9>
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (8. Aufl.). Routledge. <https://doi.org/10.4324/9781315456539>
- Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: a systematic review of randomised controlled trials in education research 1980–2016. *Educational Research*, 60 (3), 276–291. <https://doi.org/10.1080/00131881.2018.1493353>
- Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Demby, H., Jenner, L., Gregory, A., & Jenner, E. (2020). Structuring Evaluation Partnerships: Exploring Contrasts in Researcher-Practitioner Roles and Responsibilities When Implementing Randomized Experiments in Real-World Settings. *American Journal of Evaluation*, 41 (4), 531–546. <https://doi.org/10.1177/1098214020941089>
- Döring, N. (2023). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (6. Aufl.). Springer. <https://doi.org/10.1007/978-3-662-64762-2>
- Düllberg, A., Eberl, V., & Maniti, V. (2023). Die vernachlässigte Praxis? Landesinstitute als Schnittstellenakteur. In T. Diederichs & A. K. Desoye (Hrsg.), *Transfer in Pädagogik und Erziehungswissenschaft. Zwischen Wissenschaft und Praxis* (S. 106–119). Beltz.
- Farley-Ripple, E., May, H., Karpyn, A., Tilley, K., & McDonough, K. (2018). Rethinking Connections Between Research and Practice in Education: A Conceptual Framework. *Educational Researcher*, 47 (4), 235–245. <https://doi.org/10.3102/0013189x18761042>
- Gärtner, H. (2024). Nutzung von Lernstandserhebungen zur Förderung adaptiven Unterrichts. In A. Albers & N. Jude (Hrsg.), *Blickpunkt Bildungsmonitoring – Bilanzen und Perspektiven* (S. 142–157). Beltz. <https://doi.org/10.3262/978-3-7799-7843-5>
- Gärtner, A., Leising, D., & Schönbrodt, F. D. (2023). Empfehlungen zur Bewertung wissenschaftlicher Leistungen bei Berufungsverfahren in der Psychologie. *Psychologische Rundschau*, 74 (3), 166–174. <https://doi.org/10.1026/0033-3042/a000630>
- Gillon, G., McNeill, B., Scott, A., Gath, M., Macfarlane, A., & Taleni, T. (2024). Large scale implementation of effective early literacy instruction. *Frontiers in Education*, 9:1354182.
<https://doi.org/10.3389/feduc.2024.1354182>
- Gopalan, M., Rosinger, K., & Ahn, J. B. (2020). Use of Quasi-Experimental Research Designs in Education Research: Growth, Promise, and Challenges. *Review of Research in Education*, 44 (1), 218–243. <https://doi.org/10.3102/0091732x20903302>

- Hartmann, U., & Kunter, M. (2022). Mehr Praxis in der Bildungsforschung? Eine Studie zu Praxisperspektiven in Forschungsprojekten. *Bildungsforschung*, 2, 1–24. <https://doi.org/10.25656/01:25469>
- Hedges, L. V. (2018). Challenges in Building Usable Knowledge in Education. *Journal of Research on Educational Effectiveness*, 11 (1), 1–21. <https://doi.org/10.1080/19345747.2017.1375583>
- Hetmanek, A., Diery, A., Knogler, M., Schneeweiss, A., & Seidel, T. (2023). Das Clearing House Unterricht. Eine Plattform zur Stärkung der Evidenzbasierung in der Aus- und Weiterbildung von Lehrkräften. In K.-S. Besa, D. Demski, J. Gesang & J.-H. Hinzke (Hrsg.), *Evidenz- und Forschungsorientierung in Lehrer*innenbildung, Schule, Bildungspolitik und -administration. Neue Befunde zu alten Problemlagen* (S. 107–121). Springer VS. https://doi.org/10.1007/978-3-658-38377-0_6
- Holtappels, H. G. (2019). Transfer in der Schulentwicklung. Ansätze und Gelingensbedingungen aus der Perspektive von Schulentwicklungstheorie und -forschung. *DDS – Die Deutsche Schule*, 111 (3), 274–293. <https://doi.org/10.31244/dds.2019.03.03>
- Holzer, J., Grützmacher, L., Ludwig, S., Bacher, J., Dumont, H., Kampa, N., Krainer, K., Lüftenegger, M., Maaz, K., Pant, H. A., Prenzel, M., Spiel, C., & Schober, B. (2024). Bildung gemeinsam gestalten: Empfehlungen für Projekte in Kooperation zwischen Wissenschaft, Politik und Praxis. *Zeitschrift für Erziehungswissenschaft*, 28, 151–179. <https://doi.org/10.1007/s11618-024-01273-0>
- KMK (Kultusministerkonferenz). (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. LinkLuchterhand. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/Beschluesse_Veroeffentlichungen/Bildungsmonitoring_Broschuere_Endf.pdf
- KMK (Kultusministerkonferenz). (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. KMK. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf
- Leutner, D. (2013). Perspektiven pädagogischer Interventionsforschung. In E. Severing & R. Weiß (Hrsg.), *Qualitätsentwicklung in der Berufsbildungsforschung* (S. 17–28). Bertelsmann. <https://doi.org/10.3278/111-054w017>
- Lewalter, D., Diedrich, J., Goldhammer, F., Köller, O., & Reiss, K. (Hrsg.). (2023). *PISA 2022. Analyse der Bildungsergebnisse in Deutschland*. Waxmann. <https://doi.org/10.31244/9783830998488>
- Maag Merki, K. (2021). Empirische Bildungsforschung im deutschsprachigen Raum. Rückblick und Ausblick. *Swiss Journal of Educational Research*, 43 (1), 41–50. <https://doi.org/10.24452/sjer.43.1.4>
- Maaz, K., & Marx, A. (Hrsg.). (2024). *SchuMaS – Schule macht stark. Sozialraumorientierte Schul- und Unterrichtsentwicklung an Schulen in schwierigen Lagen*. Waxmann. <https://doi.org/10.31244/9783830999102>
- Pas, E. T., Ryoo, J. H., Musci, R. J., & Bradshaw, C. P. (2019). A state-wide quasi-experimental effectiveness study of the scale-up of school-wide Positive Behavioral Interventions and Supports. *Journal of School Psychology*, 73, 41–55. <https://doi.org/10.1016/j.jsp.2019.03.001>
- Prediger, S., Selter, C., Götzke, D., Hallemann, S., Holzäpfel, L., Kreuziger, A., Pant, H. A., & Rösken-Winter, B. (2024). QuaMath – Unterrichts- und Fortbildungsqualität in Mathematik entwickeln: Konzept des Zehnjahres-Programms von DZLM und KMK. *GDM-Mitteilungen*, 116 (1), 49–61. <https://qmarchiv.dzlm.de/dokumente/praxispublikation/179>
- Prenzel, M. (2019). Von SINUS zu SINUS-Transfer. *DDS – Die Deutsche Schule*, 111 (3), 331–333. <https://doi.org/10.31244/dds.2019.03.08>
- Schmidt-Hertha, B., Tervooren, A., Martini, R., & Züchner, I. (2024). *Datenreport Erziehungswissenschaft 2024*. Budrich. <https://doi.org/10.2307/jj.11786270>

- Slavin, R. E., Cheung, A., Holmes, G., Madden, N. A., & Chamberlain, A. (2013). Effects of a Data-Driven District Reform Model on State Assessment Outcomes. *American Educational Research Journal*, 50 (2), 371–396. <https://doi.org/10.3102/0002831212466909>
- Strohmer, J., & Wirtz, M. A. (2022). Theorie- und hypothesenprüfende Studiendesigns. In M. Haring, C. Rohlf, & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (2. Aufl., S. 868–879). UTB.
- Thiel, F. (2014). Evidenzbasierte Bildungspolitik – Generierung und Nutzung wissenschaftlichen Wissens. In Bundesministerium für Bildung und Forschung (BMBF) (Hrsg.), *Bildungsforschung Band 40. Bildungsforschung 2020 – Herausforderungen und Perspektiven* (S. 116–127). BMBF.
- Von Maurice, J., Auriga, R., Bayer, S., Decristan, J., Holzberger, D., Karst, K., Kuger, S., Schindler, S., & Wecker, C. (2025). *Positionspapier der Gesellschaft für Empirische Bildungsforschung (GEBF) – „Den Zugang zu Schulen und Daten aus Schulerhebungen für die Empirische Bildungsforschung verbessern“*. https://table.media/assets/briefings/bildung/positionspapier_gebf.pdf
- Weishaupt, H., & Rittberger, M. (Hrsg.). (2012). *Bildungsforschung in Deutschland – eine Situationsanalyse*. BMBF. <https://doi.org/10.25656/01:8163>.
- Zapp, M., & Powell, J. W. (2017). Moving towards Mode 2? Evidence-based policy-making and the changing conditions for educational research in Germany. *Science and Public Policy*, 44 (5), 645–655. <https://doi.org/10.1093/scipol/scw091>

Holger Gärtner, Univ.-Prof. Dr., Leiter des Arbeitsbereiches Unterrichts- und Schulevaluation an der Freien Universität Berlin.

E-Mail: holger.gaertner@fu-berlin.de, ORCID: 0000-0001-8814-9341