

Heinz Metzen

Vom Evidenz- zum Lernparadigma empirisch-pädagogischer Forschung

Zur methodischen Inklusion der Forschenden und Beforschten, der Professionsgemeinschaften und der Anwendungsinteressierten in den Forschungsprozess qua theoretischer, organisatorischer und technisch-medialer Förderung von Metareflexion des gemeinsamen Vorgehens im Prozess des Forschens. Entwurf einer integrativen Forschungsmethodik

Zusammenfassung:	2
1 Einleitung + Kontext	2
2 Exemplarische Belege für die unbefriedigende Situation der empirisch-pädagogische Forschung	3
3 Die Logik der neuen Forschungsmethodik	7
3.1 Forschungsmethodik und ForscherInnenerleben	9
3.2 Realistische Wissenschaftstheorie und Methodologie	10
3.3 Praktisch-methodische Verfahren der Forschungsprozessorganisation	13
3.3.1 <i>Zur Integration von verfügbarem und genutztem Wissen, von Forschungsziel und Erkenntnisbedarf</i>	15
3.3.2 <i>Formen reflexionsfördernder Dokumentation und Explikation des Forschungsprozesses</i>	17
Schluss	21
Literatur	22

Preprint des Artikels:

Metzen, Heinz (2016): Vom Evidenz- zum Lernparadigma empirisch-pädagogischer Forschung. Entwurf einer integrativen Forschungsmethodik. In: Hahn, Heike; Esslinger-Hinz, Ilona; Panagiotopoulou, Argyro (Hrsg.): Paradigmen und Paradigmenwechsel in der Grundschulpädagogik. Entwicklungslinien und Forschungsbefunde. Für Prof. Dr. Ursula Carle zum 65. Geburtstag. 1. Auflage. Baltmannsweiler: Schneider Hohengehren (Entwicklungslinien der Grundschulpädagogik, Band 12), S. 232–251.

Zusammenfassung:

Fortschritte bei der sozialen und pädagogischen Inklusion bedürfen auch der Unterstützung durch die empirisch-pädagogische Forschung (Abkürzung: EPF). Dass die EPF diese Unterstützung nur unzureichend leistet, hat viele Gründe. Diese Abhandlung möchte einen forschungsinternen Grund behandeln. Ich nenne ihn die in der EPF vorherrschende „*Methodologische Segregation*“. Diese findet ihren Ausdruck vor allem in der sachfremden Trennung von Forschenden und Beforschten, von verfügbarem und genutztem Wissen, von Forschungsziel und Erkenntnisbedarf, von Forschungsfrage und Forschungsmethodik, von Forschungsprozess und Forschungsbericht, von methodologisch-dokumentarischer Konvention und methodischer Objektangemessenheit, von Komplexität des Forschungsfeldes und Begrenztheit der Forschungskapazitäten und vor allem von konkurrierender Trennung der zig unterschiedlichen Forschungsansätze. Epistemisch ist die szientistische Segregation so wie die schulisch-unterrichtliche Segregation eng verbunden mit einer eingeschränkten Lernqualität.

Der Artikel belegt diese problematische Situation der EPF anhand ausgewählter methodenkritischer Reviews. Eine Lösung wird im Entwurf einer integrativen, systemisch angelegten Skizze von realistischer Wissenschaftstheorie, verschiedenen pragmatistisch-systemischen Methodologien und einigen praktisch-methodischen Verfahren der Forschungsprozessorganisation versucht. So ergibt sich eine Forschungsmethodik, die die konventionellen Vorgehensweisen nicht ersetzt, aber wirkungsvoll ergänzt. Deshalb könnte ich sie modisch als *Forschungsmethodik 2. Ordnung* charakterisiert werden. Alle ihre Ingredienzien sind nicht neu, wenn auch vielleicht nicht überall üblich. Sie sind alle praxiserprobt und ohne großen Mehraufwand selbst in kleineren Forschungsvorhaben anwendbar. Neu ist der Versuch der Überwindung konventioneller methodologischer Barrieren und, dass der Kreativität und dem Lernen im Forschungsprozess ein methodisch gestützter, größerer Spielraum gewährt wird.

1 Einleitung + Kontext

Seit der Arbeit an meiner Dissertation zur forschungsmethodischen Qualität von Evaluationsstudien in der Auswärtigen Kulturpolitik der Bundesrepublik Deutschland (Metzen 1977) bewegt mich das Problem einer validen und belastbaren Forschungsmethodik im pädagogischen Bereich. Inzwischen hat sich meine Sicht auf dieses Problem erheblich erweitert. Zur Forschungslogik im engeren methodischen Sinne haben sich Fragen nach der Wissenschaftspolitik, der Praxisrelevanz, der Wissenschaftstheorie, der Professionalisierung und der Forschungsorganisation gesellt. Das gesamte Problemfeld ist so weit gespannt und so komplex, dass die Frage nach den Ursachen des mangelnden Beitrags

der EPF zur Lösung der anstehenden Integrationsaufgaben in der Pädagogik im Rahmen eines kurzen Essays unmöglich zu beantworten ist.

Die Auswahl des hier darstellbaren Schwerpunktes hat keinen systematischen Hintergrund, sondern orientiert sich - der Thematik des Bandes entsprechend - an der dokumentierten und z. T. vom Autor miterlebten Forschungspraxis von Ursula Carle. Diese Forschungspraxis ist lebensnah, entwickelnd, persönlich, systemisch und pragmatistisch, Forschung als *Begleitforschung*, dem praktischen Problem und potenziellen Lösungen beim Entwicklungsprozess der Schulen mit den Akteuren der verschiedenen Ebenen mehrperspektivisch auf die Spur kommend. Einige methodologische und methodische Ingredienzien dieser Praxis will dieser Beitrag beschreiben.

2 Exemplarische Belege für die unbefriedigende Situation der empirisch-pädagogische Forschung

Ihren Niederschlag findet die „vorherrschende methodologische Segregation“ der empirisch-pädagogischen Forschung vor allem in der dokumentarisch, theoretisch und methodisch kaum nachvollziehbaren Struktur nahezu aller publizierten, empirisch-pädagogischen Forschungsberichte, -Essays und -Artikel. Dadurch bleibt die *Rolle der Forschenden wie der Beforschten* in diesen Dokumenten im Verborgenen. Das behindert die Forschungs- und Anwendungsgemeinschaften bei der Nutzung dieser Forschungsbefunde und bei ihrer Professionalisierung. Und es beeinträchtigt die Forschenden bei ihrem Lernen und bei ihrem Erkenntnisgewinn im Forschungsprozess sowie bei ihrer professionellen Entwicklung. Ob tatsächlich oder nur den dokumentarischen Schwächen geschuldet - es finden sich bei EPF-Dokumenten in der Regel eine unterkomplexe und deshalb gegenstandsunangemessene methodische Anlage der Forschungsvorhaben. Hierzu als Beleg einige ausgewählte, methodenkritische Reviews.

Noch eine Vorbemerkung: Das Urteil „vorherrschende methodologische Segregation“ schließt das bunte Neben- und Gegeneinander der vielerlei forschungsmethodischen „Schulen“ nachdrücklich aus. Mit Paul Feyerabend bin ich sogar der Meinung, dass ein *theoretischer und methodischer Pluralismus* dem Wissensfortschritt mehr dient als ein diesbezüglicher rigider Monismus (Feyerabend 2002, S. 15). Diese Forderung ist pragmatischer Natur und folgt der Logik, dass kluge Vielfalt besser ist als beschränkte Standardisierung. Und sie folgte einem vernichtenden Urteil, zu dem Paul Feyerabend Mitte der 1960er Jahre gelangte, nachdem er die Kohärenz von Empiristischer Wissenschaftstheorie und der diese für sich reklamierenden, praktischen Laborforschung in der Physik untersucht hatte. Weder Kopernikus, Galilei, Newton oder die modernen QuantenphysikerInnen haben sich je in ihrer praktischen Forschung an das gehalten, was wir wissenschaftstheoretisch als Empirismus bezeichnen: Das Primat des Empirischen und die Trennung von Theorie und Beobachtung. Hätten sie sich daran gehalten, was PhysikerInnen spätestens seit

Newton als forschungsbestimmend predigten, sähen wir die Erde wahrscheinlich noch als Scheibe, um die sich die Sonne wärmend dreht. Auch 50 Jahre später kann Feyerabend's Urteil ohne Abstriche auf das Verhältnis von Wissenschaftstheorie und Forschungspraxis in der Empirisch-pädagogischen Forschung übertragen werden. Den Nachweis dazu werde ich an dieser Stelle schuldig bleiben. Was ich aber im Folgenden etwas weiter auszuführen und zu begründen versuche, sind die erheblichen theoretischen und methodischen Schwächen, die die EPF kennzeichnen und für die auch das Wissenschaftsverständnis eine wichtige Rolle spielt.

Die empirisch-pädagogische Forschung hat noch keine eigene Forschungsidentität aufgebaut, folgt vielmehr weitgehend dem über die Psychologie vermittelten *Forschungsmodell der Medizin*¹. Dabei reproduziert sie auch deren Probleme, steckt das medizinische Forschungsmodell doch aktuell in einer tiefgreifenden Krise. Das jedenfalls deuten zwei Ereignisse der jüngeren Zeit an. Prof. Dr. Richard Horton FMedSci, Chefredakteur und Herausgeber der angesehensten medizinischen Zeitschrift der Welt, The Lancet veröffentlichte im Frühjahr 2015 ein erschütterndes Eingeständnis: Eine erschreckende Zahl wissenschaftlicher pharmazeutisch-medizinischer Publikationen sei bestenfalls unzuverlässig, wenn nicht völlig falsch oder sogar betrügerisch. Erst vor kurzem unterstrich er noch einmal den Ernst der Publikationsmisere in der medizinischen Forschungsliteratur, indem er die HerausgeberInnen medizinischer Forschungsorgane beschuldigte, für die „Krankheit des heutigen Verlagswesens“, für die „Verwüstungen des wissenschaftlichen Publizierens“ und für die „Verödung der Herausgeberidentität“ verantwortlich zu sein (Horton 2015; 2016 - freie Übersetzung d. A.). Ein weiteres Problem bildet die vorherrschende test-experimentelle Effektmaß-Methodologie und metaanalytische Forschungssynthese der „Evidenzbasierten Medizin“. Bereits 2003 forderte die WHO Reproductive Health Library zur Gewinnung verlässlicher Daten „Große und einfache Wirkungsstudien / Large simple trials“ (vgl. Villar / Duley 2003). Zehn Jahre später pocht das US-amerikanische Institute of Medicine der National Academy of Sciences auf dasselbe methodologische Vorgehen für die gesamte biomedizinische Forschung der USA (vgl. Grossmann / Sanders / English 2013). Damit ist der naturalistisch-medizinische Positivismus kein Vorbild mehr für die empirisch-pädagogische Forschung. Weder die anwendungsnahen test-experimentellen Effektmaßstudien noch ihre statistisch-metaanalytische

¹ Zum Beleg für die bis in die siebziger Jahre wirksame Hegemonie des medizinischen Forschungsparadigmas: Bildung und Soziales gehörten in den USA bis 1979 zum Gesundheits-Ministerium. Das Head Start Programm der Johnson-Regierung und seine Evaluation durch die Westinghouse Learning Corporation liefern ein ebenso konkretes wie krasses Beispiel für die Dominanz der MedizinerInnen in den Bildungs- und Sozialprogrammen: Die Ausgestaltung des Head Start Programms wurde stark durch MedizinerInnen dominiert. In Deutschland war die medizinische Schuleingangsuntersuchung lange Pflicht und hatte bis vor wenigen Jahren noch erheblichen Einfluss auf die pädagogischen Fragen der Zurückstellung bzw. der Förderschulüberweisung.

Synthesen erbringen in der Medizin und noch viel weniger in der Pädagogik anwendbare und zuverlässige Erkenntnisse.

Doch selbst wenn man die Konsequenzen aus diesen metawissenschaftlichen Ertragsanalysen außer Acht ließe, folgen auch heute noch viele empirisch-pädagogische Studien nicht einmal dem methodologisch angestrebten Vorgehensmuster. Im Gegenteil, die allermeisten Studien entsprechen methodisch ausnahmslos der unzureichenden Qualität, wie sie Anthony Onwuegbuzie und Larry Daniel bereits 2003 in einer Literaturübersicht zu *Analyse- und Interpretationsfehlern* in qualitativen und quantitativen Forschungsstudien aus der Pädagogik belegen. Dazu untersuchten sie rund 200 Studien zur Qualität von wissenschaftlichen Forschungsberichten. Ihr beinahe erschütternder Befund (ebd., S. 23 f):

„Eine Vielzahl von analytischen und interpretatorischen Fehlern findet sich sowohl in der quantitativen wie auch in der qualitativen Forschung. Auf der Grundlage der Häufigkeit der identifizierten Fehler muss man sich fragen, wie viel Prozent der veröffentlichten pädagogischen Forschungsergebnisse ungültig sind. Auf jeden Fall wird deutlich, dass man Studien mit äußerster Vorsicht begegnen sollte, bei deren Durchführung quantitative wie qualitative Analysen der Komplexitätsgrad des Gegenstandes unberücksichtigt blieb. Dabei bildet der Einsatz von anspruchsvollen analytischen Techniken und aufwendiger Computer-Software keinen Ersatz für die genaue Kenntnis der zugrunde liegenden Daten und die sorgfältig Prüfung aller Vorannahmen“ (freie Übersetzung d. A.).

Für einen Beitrag im „Handbuch frühkindliche Bildungsforschung“ (Stamm / Edelmann 2013) untersuchten Ursula Carle und der Autor die *Entwicklung der Evaluationsforschung* in den rund 50 Jahren ihrer öffentlichen Anwendung auf große frühpädagogische Förderprogramme. Das Positive: Die Evaluation hat sich zu einer weltweit beachtlichen Profession entwickelt. Das Negative: Heute streiten über hundert unterschiedliche Evaluierungs-Ansätze an methodischen Zugängen bzw. inhaltlichen Nutzenargumenten. Diese Evaluationsparzellen werden von großen und kleinen LehnsherrInnen (Scriven 2004: „Fiefdom Problem“) eifersüchtig verteidigt. Im frühkindlichen Bereich produzierten deren Evaluationen jedoch überwiegend Befunde, die den untersuchten Entwicklungsvorhaben mehr oder weniger Wirkungslosigkeit attestieren (Carle / Metzzen, 2013, S. 857 ff).

Nach unserer Analyse haben die Evaluations- und die frühpädagogischen Förder- bzw. Entwicklungsprogramme noch nicht die *institutionelle und professionelle Autonomie* erreicht, die zur gegenseitigen Befruchtung notwendig wäre. Der Evaluation mangelt es überwiegend noch an der Fähigkeit, sich als Teilfunktion eines Unterstützungssystems für diese Förderprogramme zu sehen, ohne dabei ihre wissenschaftlichen Standards und Forschungsinteressen aufzugeben. Auch die Frühpädagogik sieht sich bis dato noch kaum imstande, der

Evaluation auf Augenhöhe zu begegnen und sie als notwendige Entwicklungsbedingung zu verstehen und zu nutzen.

„Beide Disziplinen stehen also vor diesem letzten Schritt hin zu einem beginnenden koevolutiven Verhältnis. Dieser letzte Schritt verlangt von der wissenschaftlichen Evaluation die Überwindung überkommener naturwissenschaftlich-medizinischer Forschungsstrategien... Das Verhältnis des Förderprogramm-Evaluations-Soziotops zum gesellschaftlichen Umfeld, verdiente ebenfalls eine vertiefte Betrachtung. Schliesslich bestimmt das gesellschaftliche Interesse an beiden Disziplinen, ob ihnen weitere Entwicklungsressourcen zur Verfügung gestellt werden. Allerdings ist dieses Äußere auch aus der historischen Perspektive durchaus zwiespältig zu sehen. So kam insbesondere der Evaluation bislang in Bezug auf den Prozess der demokratischen und prosozialen Verantwortungsübernahme eine nicht selten äußerst ambivalente Rolle zu. Sie wurde sowohl als Hilfsmittel eingesetzt, um aus Erfolgen wie aus Misserfolgen der Förderprogrammentwicklung zu lernen. Sie wurde aber auch missbraucht, um das „Gelernte“ als Argument in entwicklungsfeindliche politische Entscheidungsprozesse einfließen zu lassen. Auch hiergegen würde eine vertiefte Kooperation wappnen“ (Carle / Metzen, 2013, S. 869 f).

Abgesehen von diesen lähmenden institutionellen und bildungs- bzw. sozialpolitischen Problemen der Evaluationsdisziplin, litten die meisten Studien - besser: Publikationen der Studien - unter methodischen Schwächen und unzureichender Gegenstandsangemessenheit ihrer Erhebungen. Dieser Befund lässt sich unserer Meinung nach auf viele Bereiche der empirisch-pädagogischen Forschung übertragen, wie eine Expertise derselben AutorInnen zum Wirksamkeitsnachweis des „Jahrgangsübergreifenden Lernens (JüL)“ ergab (Carle / Metzen 2014). Das für den Grundschulverband erstellte Gutachten gelangt zu einem, den methodologischen Mainstream in der Pädagogik düpierenden Befund. Die umsetzungsnahen Praxisberichte ergeben insgesamt ein umfassenderes und pädagogisch tragfähigeres Bild von den Gestaltungsbedingungen und Wirkungspotentialen des JüL als die insgesamt 70 Jahre umfassenden, mehr als hundert „quantitativen“ (testexperimentellen Effektmaß-) Studien - einschließlich des mangelhaften Effektmaßmittelungskalküls von John Hattie (ders. 2013; zur Kritik s. Carle / Metzen 2014, S. 7, 90, 93, 105, 109).

„Obwohl die schulische Praxis von der Bewährung des JüL überzeugt ist, hört die Bildungspolitik heutzutage eher auf die Stimme der testexperimentell orientierten Wirkungsforschung. Im Kapitel 5 zeigt das Studium von rund 40 Reviews (mit über 600 angegebenen, sich z.T. überschneidenden Primärquellen) sowie etwa einem Dutzend Primärstudien zum jahrgangsgemischtem Unterricht zwischen 1938 und 2013, dass das testexperimentelle Forschungsparadigma seinen Ruf als »Goldstandard« der empirischen Wirkungsforschung im Bereich des Jahrgangsübergreifenden Lernens nicht gerecht wird. Die unbefriedigenden Wirkungsaussagen zum JüL aus testexperimentellen Studien lassen sich empirisch kaum nachvoll-

ziehen und entstammen wohl überwiegend dem Wunschenken der Forschenden und ihrer Auftraggeber“ (Carle / Metzen 2014, S.126).

Viele dieser Studien beruhen auf einfachen Testleistungsvergleichen. Dazu resümierte erst jüngst ein Review mittelamerikanischer, pädagogischer PsychologInnen zur *Validitätsproblematik von Schulleistungstests*, dass hinreichende Beweise für die Gültigkeit der pädagogischen Test-Leistungsmaße entweder nicht vorliegen, nicht berichtet werden oder für eine unabhängige Überprüfung nicht zugänglich sind (s. Della-Piana / Kubo Della-Piana / Gardner 2015). So kann abschließend das Sample der hier in Betracht gezogenen fremden und eigenen methodologischen Reviews wohl als eine quantitativ wie international repräsentativ erscheinende Zahl empirisch-pädagogischer Forschungsstudien angesehen werden. Die Evidenz der methodologischen Kritikpunkte an der Empirisch-pädagogischen Forschung kann also als ausreichend belegt gelten.

Und damit sind wir wieder bei meinem unguuten Gefühl, das mich während meiner Arbeit an meiner Dissertation vor gut vierzig Jahren beschlich: „Wissenschaftstheorien“ gehören zwar zur Pflichtrhetorik jeder empirisch arbeitenden PädagogIn, aber in keiner der so nobilisierten Arbeiten schlug sich der hohe Anspruch auch nur andeutungsweise methodologisch nieder. Oder besser gesagt: In den so verzierten Berichten über die empirische Forschung fand sich keine Vorgehensweise, die wissenschaftstheoretisch nachvollziehbar begründet gewesen wäre. Im Gegenteil, die meisten Berichte empirischer Untersuchungen, die mir seit diesen Tagen begegneten, entsprachen noch nicht einmal den grundlegenden methodischen Ansprüchen der Sozialforschung (als ein Exempel für diese Standards s. Friedrichs 1999).

Der folgende methodologische Entwurf versucht eine positive Wendung der wichtigsten Kritikpunkte mit einem integrativen Impetus auf mehreren Strukturebenen. Ein kritisches Moment bleibt davon ausgenommen, die große Lücke zwischen den allerorten hoch gehaltenen Methodologien / Wissenschaftstheorien und der beobachteten wie der dokumentierten Forschungspraxis.

3 Die Logik der neuen Forschungsmethodik

Zur hinreichenden Charakterisierung eines neuen methodologischen Paradigmas sind in mindestens vier Kernbereichen Klärungen notwendig: (1) Die wissenschaftsphilosophischen Annahmen und Haltungen, (2) die neue Untersuchungslogik (Methodologie), (3) die zusätzlichen Methoden, Techniken und Leitlinien für die Forschungspraxis sowie (4) die sozialen und politischen Verpflichtungen der Forschenden (nach Green 2006; Onwuegbuzie / Johnson / Collins 2009). Dass im Folgenden der vierte, der sozialpolitische Bereich außer Acht gelassen wird, ist der Ökonomie dieses Beitrags geschuldet. Der einleitende Bezug zur Integrationspädagogik mag als Andeutung der Ausrichtung dieses fehlenden Aspektes erst einmal genügen.

Die hier unterstellte *Ertragsarmut*² der empirisch-pädagogischen Forschung für die Integrationspädagogik und für viele andere pädagogische Arbeitsfelder hat - wie schon angedeutet - neben anderen v. a. methodologische und methodisch-technische Ursachen. Darüber hinaus stehen die „Soft Sciences“, zu denen die Pädagogik gehört, hinsichtlich ihres Forschungsgegenstandes vor zusätzlichen „Hard Problems“ - wie Heinz von Foerster immer wieder sprachspielerisch betonte. Ihr Gegenstand konstituiert sich nicht nur aus naturwissenschaftlichen Faktoren, sondern zudem aus medizinischen, sozialen und kognitiven Gegebenheiten plus deren Wechselwirkungen. Dass zur Erforschung dieses meta-disziplinären Komplexes die einfachen Werkzeuge der „harten“ Natur- und Technikwissenschaften zu wenig griffig sind, liegt auf der Hand. Diese leicht nachvollziehbare Konflikt wird bis dato vom test-experimentellen Effektmaß-Mainstream weitgehend ignoriert. Was also bietet eine Lösung des „harten“ methodologischen Problems der „weichen“ Pädagogikforschung? Wie angekündigt soll eine „integrative *Forschungsmethodik*“ diese Lösung bieten. Deren Kernmoment bildet die Integration verschiedener Ebenen und Phasen der wissenschaftlichen Kooperation und Kommunikation. Und dies sind ihre Hauptingredienzen (bereits angedeutet, skizziert und z. T. ausgearbeitet in Carle 2000):

- Eine „realistische Wissenschaftsauffassung“ mit den methodischen Konsequenzen
- die Explikation des ForscherInnenerlebens und Handelns (Protokollreflexion, Wirkungsmodell, Hypothesenentwicklung, Entscheidungsmatrizen)
- die Integration der Beforschten in den Forschungsprozess
- die Notwendigkeit von Feldkompetenz auf Seiten der Forschenden und von Forschungskompetenz auf Seiten der „Beforschten“
- die Verbesserung der Forschungsprozessorganisation
- die Anwendung von „Komplexmethoden“ als Erhebungsinstrumente
- die intensivere und detailliertere, der Reflexion und Überprüfung zugängliche Forschungsprozessdokumentation und
- die Trennung dieser Forschungsdokumentation von der folgenden Berichtspublikation³

² Ertragsarmut weniger die Forschenden, denn Ertrag gibt es für sie genug. Denn die Scientific Community goutiert die Forschung wesentlich über die nachgewiesene Zugehörigkeit zu der jeweiligen Methodenrichtung. Da sich die „Schulen“ immer mehr aufsplitten, darf man sich nur dort verorten und entsprechend eng binden. Wissenschaft schafft nicht mehr Wissen, sondern eine immer bessere Abgrenzung von methodisch illegitimen Schulen. Das forschende Handeln ist somit defensiv, nicht expansiv orientiert. Der Ertrag wird über die Abwehr von Rufschäden und Ausschlüssen ausgezahlt.

³ Der sich hier anbietende Vorwurf der erneuten „Segregation“ stimmt insofern nicht, als Ideen zur Publikation sehr wohl forschungsprozessbegleitend gesammelt werden sollten. Die Separation der Berichts-Publikation ergibt sich als eine Reaktion auf die vorherrschende Vermischung von Publikation und Forschungsprozessdokumentation, worunter beide leiden; zudem folgt diese Trennung der Forschungsschritte-Logik.

3.1 Forschungsmethodik und ForscherInnenerleben

„Forschungsmethodik 2.0“ lehnt sich an die Bezeichnungsform des sozialen Konstruktivismus an (Foester 1993, S. 60 ff). Danach beobachten die ForscherInnen nicht nur den Forschungsgegenstand im Forschungsfeld, sondern auch sich, d. h. ihr Beobachten, Reflektieren und Handeln im Forschungsprozess. Dieser Sicht liegt die Annahme zugrunde, dass die Trennung zwischen den BeobachterInnen / ForscherInnen und den Beforschten / dem Forschungsgegenstand im sozialen Feld praktisch unmöglich ist. Im Gegensatz zu rein stofflich-energetischen Naturwissenschaften „äußert“ der Gegenstand der Pädagogik seine Reaktionen nicht als unpersönliche, „objektive“ Evidenzen, die er zwischen sich und den Forschenden auf einen imaginären Labortisch legt. Vielmehr manifestiert sich die pädagogische Evidenz seiner Interaktionen mit uns und die unseres Handelns mit dem Gegenstand vor allem in Form unseres Erlebens (bewusst und unbewusst). Diese inter- und intrapersonale Manifestation der sozialen Reaktionen auf die Forschungs-Intervention ist uns nur über die Explikation unseres Erlebens und unseres Handelns inklusive der ihm zugrunde liegenden Kenntnisse und Überzeugungen sowie über die Explikation zugänglich. Die Daten der Verhaltensprotokolle dienen dabei primär als Erinnerungstützen und sekundär als kategorial-mengenmäßige Reflexionshilfe. Die integrative Forschungsmethodik geht diese *Explikation des Erlebens der ForscherInnengemeinschaft*⁴ methodisch an. Sie sucht nach Werkzeugen zur Unterstützung dieser Selbstreflexion und nach Möglichkeiten, an dieser Selbstreflexion auch externe Fachkundige und die Forschungsstakeholder teilnehmen zu lassen.

Dahingegen betrachtet die empiristisch-experimentelle Forschungsmethodik 1.0 diese Selbstreflexion in der offiziellen Kommunikation wenn schon nicht als störend, dann allenfalls als zierendes Beiwerk. Im internen Forschungsdiskurs spielt die prozessbegleitende Selbstreflexion auch in der empiristischen (naturalistischen, positivistischen) Forschungspraxis nach meiner Erfahrung durchaus eine nicht unbedeutende Rolle. Sie ist der Quell von kreativen Eingebungen und schöpferischen Hypothesen, die den Erkenntnisprozess oft entscheidend weiter bringen. In der offiziellen Publikationspraxis werden diese Einfälle aber hinter ähnlichen Ideen von Fachautoritäten versteckt oder in die, den Interpretationen folgenden Schlussfolgerungen eingebettet. Die „Forschungsmethodik 1.0“ folgt dem *szientistischen Paradigma des Trennens, Teilens, Spaltens*. Immerhin führte in den angewandten medizinischen und technischen Wissenschaften die analytische Suche nach den „konstitutiven Merkmalen und Beziehungen ... [einer] unbestimmten Situation“ (Dewey 2002-1938,

⁴ Nicht zu verwechseln mit dem Ansatz der „Rekonstruktiven Forschung“, denn die Rekonstruktion bringt nicht automatisch Ergebnisse, die zu einer Lösung des Problems, zur Weiterentwicklung z. B. einer Schule beitragen. Dazu muss eine gemeinsame Zielrichtung zwischen Forschenden und den die Schule Entwickelnden vorhanden sein und gemeinsam um die Interpretation der Ergebnisse mit Blick auf die bessere Zielerreichung gerungen werden.

S. 131) zu den ersten großen Fortschritten. In den Sozialwissenschaften blieb ein vergleichbarer Erfolg dieses Vorgehens von Anfang an aus. In der empirisch orientierten pädagogischen Forschung findet die „Methodologische Segregation“ ihren Ausdruck vor allem in der sachfremden Trennung von Forschenden und Beforschten, von verfügbarem und genutztem Wissen, von Forschungsziel und Erkenntnisbedarf, von Forschungsfrage und Forschungsmethodik, von Forschungsprozess und Forschungsbericht, von methodologisch-dokumentarischer Konvention und methodischer Objektangemessenheit, von der Komplexität des Forschungsfeldes und der Begrenztheit der Forschungskapazitäten.

Die lernorientierte, integrative Forschungsmethodik beendet dieses Versteckspiel, bekennt sich offen zur forscherschen Intuition und schafft zu ihrer Förderung qua Explikation ein solides Methodengerüst. Durch diese „*Systemische Forschungsprozessintegration*“ bildet sie das logische Gegenteil zur „Methodologischen Segregation“. Eine pädagogische Forschungsmethodik sollte sich allerdings nicht der analytischen Vorteile der „Methodologischen Segregation“ entledigen, sondern diese mit dem integrativen Potential des systemischen Ansatzes verbinden. Dieser Spagat wird im Entwurf einer integrativen, systemisch angelegten Skizze von realistischer Wissenschaftstheorie, verschiedenen pragmatistisch-systemischen Methodologien und einigen praktisch-methodischen Verfahren der Forschungsprozessorganisation versucht. So ergibt sich eine Forschungsmethodik, die die konventionellen Vorgehensweisen nicht ersetzt, sondern gegenstandsangemessen ergänzt. Die Forschungsmethodik 1.0 ist damit also nicht überholt, sondern liegt der selbstreflexiven Erweiterung der Forschungsmethodik 2.0 zugrunde.

3.2 Realistische Wissenschaftstheorie und Methodologie

Eine den realen Gestaltungs- und Entwicklungsprozessen der Pädagogik gerecht werdende Forschung ergibt sich allerdings nicht bloß aus der additiven Ergänzung des naturalistischen und positivistischen Kategorisierens und Quantifizierens um die methodologische und empirische Selbstreflexion. Entscheidender ist die Aufhebung des Widerspruchs zwischen Gegenstandskomplexität des pädagogischen Forschungsfeldes und *selbst gewählter Begrenzung der eingesetzten Forschungskapazitäten*. Ein wesentlicher Punkt dieser Selbstbeschränkung bildet die Reduzierung des Forschungsfeldes auf pädagogikferne Laborsituationen bzw. auf prozessferne, punktuelle Beobachtung, Befragung und Testung. Dahinter steht u. a. der Irrglaube an die wahrheitsfördernde Trennung von Forschenden und Beforschten. Ursula Carle steigt dagegen in ihrer Schulbegleitforschung als ganze Person(engruppe) in das pädagogische Entwicklungsgeschehen ein⁵. Im Vordergrund steht das konkrete Entwicklungspotential der Beteiligten, die Förderung der Qualität des eigenen Umgangs mit

⁵ Einen Überblick einschließlich der Forschungsberichte bietet die Forschungs-Seite auf ihrer Webseite der Universität Bremen (Carle o. J.).

Entwicklungsaufgaben der Individuen, Kollektive, Schulen und Schulsysteme. Daran vor allem bemessen sich ihre Unterstützungsangebote und ihre Publikationen.

Damit übernimmt *Schulbegleitforschung* immer auch Verantwortung für das Gelingen des Entwicklungsprozesses (vgl. z.B. Carle / Berthold 2004, S. 16f; Carle / Metzen 2006; Carle / Samuel 2007; Carle / Metzen 2009). Ohne dies hier näher ausführen zu können, bildet die Übernahme dieser „kooperativen Verantwortung“ (z. Begriff s. Nida-Rümelin 2011, Kap. VIII Kooperative Verantwortung, § 48, S. 123 ff) das Gegenstück zu den sozialen Aspekten der zientistischen Segregation. Nida-Rümelin definiert die kooperative Praxis als Verantwortungserweiterung:

„Kooperative Praxis kann nur dann gelingen, wenn die beteiligten Individuen ihren egozentrischen Standpunkt überwinden. Sie müssen ihr eigenes Handeln als konstitutiven Teil einer kollektiven, kooperativen Praxis ansehen. Die leitenden Intentionen sind dann auf das Gelingen dieser kooperativen Praxis und nicht auf die Folgen des eigenen Handelns gerichtet... Kooperative Praxis lässt sich begrifflich nicht adäquat erfassen, wenn man einem radikalen Individualismus verhaftet bleibt. Dies erklärt das Unvermögen der rationalen Entscheidungstheorie einschließlich ihres spieltheoretischen Zweiges, kooperative Praxis angemessen zu beschreiben und überzeugende Kriterien für sie zu entwickeln“ (ebd., S. 123).

Diese Haltung und das Prozedere ihrer gelingenden Forschungspraxis bilden die Grundlage für die Entwicklung des hier diskutierten Methodikkonzeptes. Es gibt einige methodologische Ansätze in der Sozialforschung, die diese Vorgehensweise prinzipiell teilen und wissenschaftstheoretisch begründen (Aktionsforschung, Entwickelnde Evaluation / Developmental Evaluation, Soziale Wirkungsforschung / Social Impact Assessment, Design experiments / Design research / Design-study research / Design-based research, Realistische Evaluation)⁶. Der Sozialforschungs-Methodologe Ray Pawson und der Sozialforscher Nick Tilley entwarfen erstmals 1997 die Grundzüge einer „Realistischen Evaluationsmethodik“ und setzten sich dabei naturgemäß intensiv mit den vorherrschenden, test-experimentellen Effektmaß-Methodologien auseinander (Pawson / Tilley 2011-1997, insbes. Kap. 3, S. 55 ff). 2013 publizierte Ray Pawson darauf aufbauend den Entwurf einer Evaluations-Wissenschaft, „The science of evaluation. A realist manifesto“ (Pawson 2013), in der er im Kapitel 1 die wissenschaftstheoretischen Grundlagen der „Realistischen Evaluation“ ausführt (ebd., S. 3 ff). Gemeinsamer Nenner und Quintessenz dieser „realistischen“ Ansätze: Wenn Sozialforschung die konstitutiven Merkmale und Be-

⁶ Literaturverweise: Aktionsforschung (Fricke 2014, S. 2013 ff); Entwickelnde Evaluation / Developmental Evaluation (Patton 2010); Soziale Wirkungsforschung / Social Impact Assessment (Arce-Gomez / Donovan / Bedgood 2015, S. 85 ff); Design experiments (Collins / Joseph / Bielaczyc 2004); Realistische Evaluation (Pawson 2013, S. 3 ff; Pawson / Tilley 2011, S. 55 ff). Alle Forschungsansätze, die hier noch fehlen, bitte ich um einen entsprechenden Hinweis.

ziehungen komplexer sozialer Entwicklungsprozesse ergründen will, muss sie sich den generativen, produktiven gestalterischen Momenten dieser Prozesse im statu nascendi zuwenden und eine explizierende Darstellung der persönlich zugänglichen Merkmale und Beziehungen versuchen. Dazu gehören sowohl die unterstützende und dokumentierende Beteiligung der Forschenden am Entwicklungsprozess wie auch die informierende und dokumentierende Beteiligung der EntwicklerInnen am Forschungsprozess.

Aus Sicht der Realistischen Wissenschaftstheorie á la Pawson besteht aber nicht nur die Notwendigkeit des Dabeiseins und Mitgestaltens pädagogischer Entwicklungsprozesse zur Gewinnung grundlegender und gestaltungssensitiver Merkmale und Beziehungen über diese Entwicklungsprozesse. Ein erkenntnisreiches Dabeisein und ein fruchtbares Mitgestalten der Forschenden setzen voraus, dass diese die zu entwickelnde Praxis nicht nur aus Büchern, sondern aus eigener Gestaltungserfahrung hinreichend „beherrschen“. Einer der Stichwortgeber der modernen, realistischen Wissenschaftstheorie, Donald T. Campbell (1916-1996) äußerte zu dieser Notwendigkeit dieses „qualitativen“ Wissens über das Untersuchungsfeld in einem Essay zur Frage, „Can We Be Scientific in Applied Social Science?“ (Campbell 1988, S. 323 - freie Übers. d. A.):

„Der dritte Punkt ist der Irrglaube, dass quantitative Daten qualitative ersetzen können. Stattdessen bildet qualitatives Wissen eine unbedingt notwendige Voraussetzung für die Quantifizierung in jeder Wissenschaft. Ohne Kompetenz auf qualitativer Ebene bleibt ein Computer-Ausdruck irreführend oder bedeutungslos. Es gelang uns nicht in unseren Überlegungen zu den Programmevaluationsmethoden, die Notwendigkeit eines ausreichenden qualitativen Kontextwissens zu betonen. Ein Beispiel dafür bildet die häufige Trennung von Datensammlung, Datenanalyse und Programmdurchführung... Diese führte zu einer Leichtgläubigkeit gegenüber den Zahlen aus dem Computer zusammen mit einer totalen Ahnungslosigkeit der Analytiker davon, was tatsächlich bei der Umsetzung des Programms und in der Testsituationen vor sich ging. Um rivalisierende Hypothesen begründet verwerfen zu können, benötigen wir eine situationsspezifische Felderfahrung [orig.: Wisdom]. Das Fehlen dieses Wissens (ob wir es Ethnographie oder Programmgeschichte oder Gerüchte nennen) macht uns unfähig, Programmwirkungen kompetent einzuschätzen und führt uns zu Schlussfolgerungen, die nicht nur schlicht falsch sind, sondern sich oft sogar als falsch in einer sozial destruktiven Weise herausstellen.“

Allerdings wächst mit der Größe der Vorhaben die Zahl und Unterschiedlichkeit der Praxisfelder so sehr, dass unterschiedlich qualifizierte und praxiserfahrene Forschende notwendig sind, um die gesamte Bandbreite der Unterstützungs-, Dokumentations- und Erhebungsbereiche hinreichend abdecken zu können. Eines der gelungenen Beispiele hierfür - wenn auch nicht für ein Entwicklungsprojekt, dafür aber für eine grundlegende und umfassende Längsschnittstudie - bildet das „Effective pre-school, primary and secondary education (EPPSE 3-16) project“ (vgl. Sylva et al. 2014). Darin zeigt sich, dass die

Ressourcenfrage keine nachrangige Rolle für relevante Forschung in der Pädagogik spielt. Insgesamt macht also die Segregation von „Forschung“ und „Praxis“, von Untersuchungs- und Entwicklungslogik in der Pädagogik wie in allen anderen, entwicklungsdynamischen Forschungsfeldern weder für die Wissenschaft noch für die praktischen Forschungsfelder einen Sinn. Im Gegenteil, ihre Überwindung verspricht einen deutlichen Mehrwert sowohl hinsichtlich der Entwicklungsdynamik als auch hinsichtlich des Forschungsertrags (Peter 2016, S. i ff).

Viel tiefer als die Feldkompetenz der Forschenden geht in jedem Fall das „*Wissen und Können der Beforschten*“. Sie müssen also in den Forschungsprozess eingebunden werden, als FeldexpertInnen ebenso wie als Ko-ForscherInnen. Das gilt nicht nur für Erziehungswissenschaftliche Forschung. So erwies es sich bei einer Arbeitsbelastungsuntersuchung als notwendig, ausgewiesene Arbeitsplatzkenner für die Mitwirkung bei der Arbeitsplatzbewertung zu qualifizieren und mit ihnen in einem dialogischen Verfahren gemeinsame Skalenergebnisse herausgearbeitet (Hoyos / Metzen 1982, S. 57 ff)⁷. Ansätze, die PraxisexpertInnen in die Forschung einbeziehen gehen in Richtung eines Sozialforschungsverständnisses, das den Forschungsprozess als wissensintensiven, d.h. wissensorganisierenden und wissenserweiternden gegenseitigen Dienstleistungsprozess in einem Entwicklungsvorhaben sieht. Beide Seiten profitieren: WissenschaftlerInnen und die kooperierenden Akteure aus dem Praxisfeld erlangen einen sehr viel tieferen und angemesseneren Einblick und damit die Basis, um erziehungswissenschaftliche Theorien und die Arbeit im Praxisfeld weiterzuentwickeln.

3.3 Praktisch-methodische Verfahren der Forschungsprozessorganisation

Pädagogische Forschung als wissensorganisierender und wissenserweiternder kooperativer *Dienstleistungsprozess in pädagogischen Entwicklungsvorhaben* bedarf einer intensiven Vorbereitung, Planung, Organisation, Erhebungstechnik, Auswertung, Dokumentation und Publikation. Der notwendige Mehraufwand ist nicht nur eine methodologische Konsequenz, sondern ergibt sich auch aus der Verantwortung der Wissenschaft gegenüber den Beteiligten des Forschungsfeldes und dessen mittelbaren Stakeholdern (zur „realistischen“ Definition wissenschaftlicher Verantwortung siehe das Kapitel XI in Nida-Rümelin 2011, S. 157 ff).

Dieser *planerische und organisatorische Mehraufwand* kann nicht alleine von den einzelnen MitarbeiterInnen aus den beiden beteiligten Bereichen – Wissenschaft und Praxisfeld - erbracht werden. Dreierlei ist zu seiner Deckung notwendig. Erstens bedarf es der Unterstützung durch die projektbezogene und projektinteressierte ForscherInnengemeinschaft, denn ohne Anerkennung des beschriebenen Ansatzes und ohne Austausch über diese Forschung, kann sie

⁷ Die Entwicklung dieses Verfahrens erfolgte unter Leitung des Autors Ende der 1970er Jahre.

sich kaum weiterentwickeln und erhält mangels Einfluss nicht die nötigen Forschungsgelder. Zweitens muss bei einem Entwicklungsvorhaben die organisationale, instrumentelle und finanzielle Durchführungskapazität auf Seiten des Projektträgers hinreichen (das Fachwort dazu: „Evaluability Assessment“ - vgl. Thurston / Potvin 2003, S. 453 ff). Aus der Kalkulation beider Kapazitäten folgt drittens die Eingrenzung / Ausweitung des Forschungsvorhabens, so dass die gegebenen Forschungskapazitäten dafür ausreichen, die interne Kohärenz, methodische Fundierung und praktische Relevanz der wissenschaftlichen Arbeit zu garantieren.

Die wissenschaftliche Verantwortung hat noch eine dritte, wenig diskutierte Seite: Die *Verantwortung der WissenschaftlerIn für sich selbst* und für das sie tragende Forschungskollektiv. In englischsprachigen Publikationen finden sich dazu seit wenigen Jahren Versicherungen zum Interessenkonflikt in den Forschungs-Abstracts („Conflict of interest statement / Conflict of interest disclosure“). Hierzu ist zu sagen, dass in einer konkurrenten Forschungslandschaft die Explikation der Forschungsmotive und Forschungsverläufe nicht automatisch publiziert werden müssen. Für das persönliche und kollektive Lernen und Erkennen bildet die umfassende Explikation und Dokumentation aber eine notwendige Voraussetzung. Für das Überleben und Vorwärtkommen in der Forschungsgemeinschaft sollte ihre Veröffentlichung dennoch vorsichtig gehandhabt werden und die Explikationen und Dokumentationen wohltdisiert einfließen.

Die im Folgenden aufgeführten Techniken und Methoden dienen vor allem zwei Aufgaben, der *Fehlerminimierung* (hinsichtlich Häufigkeit und Folgen) und dem *Lernen aus Fehlern*. Mit der Komplexität des Gegenstandes wachsen die Fehlermöglichkeiten. Der Mehraufwand für die Vorbereitung, Planung, Organisation, Erhebungstechnik, Auswertung, Dokumentation und Publikation pädagogischer Untersuchungen kann das Fehlerpotential zwar erheblich mindern, verursacht aber auch seinerseits zusätzliche Sekundärfehler. Also gilt es, mit Forschungsfehlern leben zu lernen, besser noch: aus ihnen Lerngelegenheiten zu schaffen. Dabei helfen zwei Techniken: Die eine besteht in der vorbereitenden und begleitenden *Explikation und Dokumentation*, d. h. in der Vorbereitung, Planung, Organisation, Erhebungstechnik, Auswertung und Publikation. Die zweite besteht in der *Aufteilung der Arbeiten in kleine, beherrschbare Schritte*. Metaphorisch gesprochen bleiben dadurch „Fehlritte“ ohne große Folgen und bieten die Gelegenheit „aus Fehlern zu lernen“ - „Work-based learning“ für Forschende.

Komplexere Problemstellungen und Untersuchungsvorhaben erhöhen aber nicht nur die Zahl möglicher Fehler, sondern auch die Gefahr, keine Ergebnisse zu erhalten. Denn immer noch gilt, „Poor methods get results“, einfache Untersuchungen und Fragestellungen lassen sich einfach beantworten, gleichgültig ob richtig oder falsch, egal ob gegenstandsangemessen oder nicht. Und leider gilt auch immer noch, keine Ergebnisse sind schlechte Ergebnisse und werden den ForscherInnen als Versagen angekreidet. Kurz: Wer sich auf das Abenteuer

realistischer Forschungsmethodologie und -methodik einlässt, gewinnt zwar ein Vielfaches an Einsichten, geht aber auch ein erhebliches Risiko des karriereschädlichen Konventionsbruchs ein.

3.3.1 Zur Integration von verfügbarem und genutztem Wissen, von Forschungsziel und Erkenntnisbedarf

Zurück zum „Abenteuer realistischer Forschungsmethodologie“. Welcher praktischen Logik folgen empirisch-pädagogische Untersuchungen? Neben der wissenschaftstheoretischen Strukturlogik besitzt der Forschungsprozess auch eine chronologische systemische Entwicklungslogik aus schlüssig aufeinander folgenden Arbeitsschritten / Aufgabenpaketen. Diese sind in der folgenden Grafik, „Hauptglieder der Untersuchungs-Lernkette“ als Glieder einer Kette dargestellt.



Abbildung 1: Chronologische Gliederung des Forschungsprozesses in Arbeitsschritte

Jeder Arbeitsschritt folgt einem vorausgehenden und bildet seinerseits wieder die Voraussetzung für den nachfolgenden. Nach dieser *Kettenlogik* ist die Gesamt-Qualität des Forschungsprozesses abhängig von der Qualität des „schwächsten“ Gliedes / Arbeitsschrittes. Jeder Einzelschritt kann nur das Qualitätsniveau des vorausgegangenen Schrittes erreichen. Das heißt nicht, dass die vielen Einzelarbeiten eines Einzelschrittes qualitativ durchaus stark differieren können. So kann die Qualität der Literaturrecherche und -dokumentation durchaus sehr gut sein, obwohl die Gesamtqualität der „Wissensaufbereitung“ (umgangssprachlich der „*Theorieteil*“), entweder in Abhängigkeit von der Qua-

lität der vorausgegangenen Problemlösung, Felderkundung und Forschungsfragestellung oder aber durch unzureichende Erschließung etwa des vorhandenen Praxiswissens ungenügend ausfällt. Der Nachteil der Kettenmetapher liegt in der Unterschlagung der Rekursivität von Arbeitsschritten im Forschungsprozess. Natürlich geht man vor und zurück und erreicht dabei auf jedem Teilschritt ein insgesamt höheres Niveau.

„Es gibt nichts Praktischeres als eine gute Theorie“ wird Kurt Lewin und zahlreichen anderen Geistesgrößen in den Mund gelegt⁸. Wie auch immer, der Satz wendet sich gegen die in der empirisch-pädagogischen Forschung vorherrschende Trennung von „Theorie“ und „Praxis“. Der diesbezügliche Vorteil des Kettenbildes besteht darin, dass es die Unsinnigkeit dieser Trennung veranschaulicht. Das *Praktische der Theorie* - hier als Wissen über das Feld, die Problemstruktur, die Forschungsmethodik - äußert sich unmittelbar in der Qualität der Untersuchungsfragen, der Erhebungsinstrumente, schließlich der Befunde und ihres Beitrags zur Problemlösung. Kehrt man den Satz um, „Es gibt nichts Theoretischeres als eine gute Praxis“, wird allerdings deutlich, wie pejorativ „theoretisch“ im Alltag gebraucht wird. Da lebensweltlicher und wissenschaftlicher Alltag kaum zu trennen sind, würde der Pejorativ erklären, warum das Niveau der „Theorieteile“ empirisch-pädagogischer Arbeiten häufig so niedrig ist, warum sie die Problemstellung verfehlen und stattdessen dem eklektizistischen Name-dropping und einer übersteigerten Fremdsprachlichkeit frönen. Auch wenn es bei Forschungs-Qualifikantinnen eine Tendenz zum „Praktischen“, zum Positivistischen gibt, so erklärt sich das allgemeine Theoriedefizit der EPF eher aus der Trennung von „Theorie“ und „Praxis“. Einen deutlichen Beleg dafür liefern die beiden Reviews von Carle und Metzen zur Evaluation in der frühkindlichen Bildung und zur Wirksamkeit jahrgangsübergreifenden Lernens (dies., 2013; 2014).

Erfolgsbedingung gelingender „Theorieteile“ in der EPF sind zum einen die hinreichende Vorklärung der Problemstellung und Aufklärung des Forschungsfeldes sowie die *Integration von verfügbarem und genutztem Wissen aus Praxis und Literatur*. Im Vordergrund der Wissensaufbereitung für das Forschungsvorhaben muss der von der Problemstellung und von den Forschungskapazitäten sich ableitende Erkenntnisbedarf sein. Im Vordergrund der Theorieaufgabe steht die Frage: Was muss ich wissen? Nicht: Wen muss ich zitieren?

Ein untrügliches Zeichen für Theoriedefizite in empirischen Untersuchungen ist das Fehlen bzw. die unklare Ableitung der *Untersuchungshypothesen*. Hierbei stoßen wir auf ein weiteres grundlegendes Problem der EPF, diesmal das Fehlen einer Trennung von Forschungsdokumentation und Berichtspublikation. Der schrittweise, rekursive und spiralförmige Prozess wissenschaftlichen Forschens lässt sich dokumentarisch nicht in das konventionalistische Korsett von Abstracts, Artikeln und Forschungsberichten pressen. Hier hilft nur die *Tren-*

⁸ Eine aktuelle Deutung bietet Ray Pawson (ders. 2003, "Nothing as Practical as a Good Theory")

nung von Forschungsdokumentation und Berichtspublikation. In den Technik- und Naturwissenschaften sind „Labortagebücher“ üblich und werden von der interessierten Fachleserschaft z. T. den Ergebnispublikationen in renommierten, englischsprachigen Journals vorgezogen. Diese Tradition des Forschungstagebuches sollte von der EPF aufgegriffen und mit den Mitteln der digitalen Medien nutzerfreundlich aufbereitet werden. Nur auf der Basis eines kontinuierlichen Forschungstagebuches können spontane Eingebungen, offene Fragen, beeindruckende Erlebnisse und vertiefte Reflexionen ohne abschließende Gültigkeitsbehauptung aufgehoben sowie der tatsächliche Verlauf der Erkenntnisgewinnung festgehalten und der späteren Reflexion zur Verfügung gestellt werden.

3.3.2 Formen reflexionsfördernder Dokumentation und Explikation des Forschungsprozesses

Je nach Forschungsumfang ist zusätzlich zum Forschungstagebuch auch die Anlage eines Organisationsdokuments notwendig, in das die Projektorganisation samt Mitwirkenden plus positive wie negative Verlaufsabweichungen notiert werden. Wie aber kommt die ForscherIn von der bunten Fülle des Tagebuches, des Organisationsdokuments zur distinkten Form des Ergebnisberichtes? Die Lösung liegt im Anlegen von separaten Themen-, Begriffs- und Literaturdateien, die letztlich auf die Verwendung im Abschlussbericht zielen, aber je nach Forschungsverlauf aktualisierbar sind, ohne an die Berichtslogik eines Artikels oder Buches gebunden zu sein.

Zurück zur chronologischen Gliederung des Forschungsprozesses in Arbeitsschritte. Die lineare Chronologie entspricht nicht der Epistemologie, dem Lernen im Forschungsprozess. Trotzdem bleiben beide Prozesse verbunden und zwar in unserer *prozessbegleitenden Reflexion*. Diese stützt sich neben dem Nachdenken beim Protokollieren (Tagebuch) auf drei weitere Instrumente / Techniken: Der Hypothesenentwicklung per Entscheidungsmatrix, der Schnittstellenentwicklung zwischen den Forschungsschritten und der Wirkungsmodellexplikation.

Alles Handeln hat Gründe und diese sind von tiefer liegenden Überzeugungen geleitet. Forschen heißt, die für den einzelnen Erkenntnisschritt relevanten Überzeugungen herauszuarbeiten, zu explizieren. Forschungsbezogene Überzeugungen nennen wir „*Hypothesen*“. Sie finden sich prinzipiell an allen Schnittstellen zwischen den Forschungsschritten, also zwischen Problemstellung und Formulierung der Forschungsfragen, zwischen diesen und der Entwicklung von z. B. Beobachtungskategorien für das Erhebungsinstrument usw. In der Literatur werden in der Regel nur die Arbeitshypothesen an der Schnittstelle von Problemstellung und Forschungsfragen sowie die Ergebnis- bzw. Auswertungshypothesen an der Schnittstelle von Untersuchungsfragen und Erhebungsinstrumenten angesprochen. Warum brauchen wir „Auswertungshypothesen“ als Explikation der Beziehung zwischen Forschungsfragen und Be-

obachtungsmerkmalen? Der Grund liegt im Versuch, ein logisches Dilemma pragmatisch zu lösen. Das Dilemma: Die Beobachtungssituation liefert keine direkten Antworten auf unsere Forschungsfragen, sondern nur Deskriptoren / Indikatoren eines mehr oder weniger erklärbaren Geschehens (Verhaltensverläufe / Ereignisse / Kontexte...). Wie lassen sich dennoch Antworten auf unsere Fragen durch die Beobachtung des typischen Geschehens gewinnen? Als Bindeglied zwischen dem beobachtbaren Geschehen und unseren Forschungsfragen nutzen wir unser theoretisches und praktisches Wissen. Darauf bauen wir Ergebniserwartungen / Hypothesen, also begründete Annahmen über Auftretens- und Intensitätswahrscheinlichkeiten der beobachtbaren Geschehensdetails auf. Dies ist ein kreativer Akt.

Hier entsteht also neues Wissen, jedenfalls für uns, das wir dann empirisch überprüfen. Die „Auswertungshypothesen“ spiegeln unsere Annahmen, unser Wissen hinsichtlich des erwartbaren Verlaufs der Beobachtungssituation. Je besser, professioneller, wissenschaftlicher die Qualität dieses, den Hypothesen zugrunde liegenden Wissens ist, umso ergiebiger kann die Beobachtung gelingen. Wie gewinnen wir die Auswertungshypothesen? Dazu müssen wir aus der nahezu unzähligen Vielzahl von Geschehensparametern die zur Beantwortung unserer Forschungsfragen wichtigsten und gleichzeitig real beobachtbaren Geschehensmerkmale auswählen. Hierfür benutzen wir eine *Entscheidungsmatrix*, die die Beziehung zwischen diesen wichtigen Forschungsfragen und den realistisch beobachtbaren Geschehensmerkmalen untersucht.

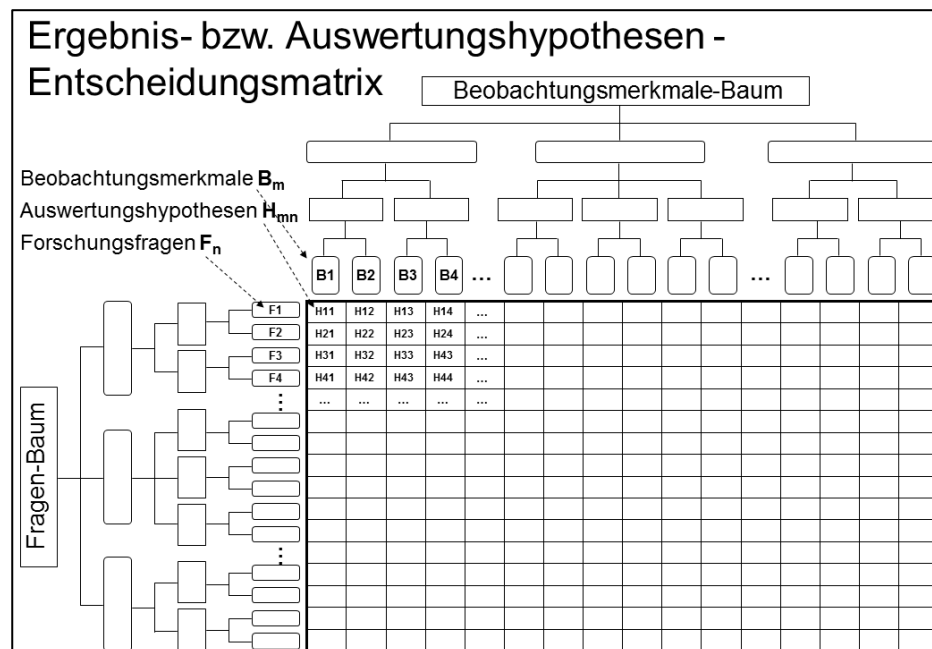


Abbildung 2: Entwicklung der Ergebnishypothesen mit Hilfe einer Entscheidungsmatrix

In der o. a. Entscheidungsmatrix wird die Gegenüberstellung von Untersuchungsfragen mit Beobachtungsmerkmalen zur Auswahl der für die Gestaltung des Erhebungsinstrumentes relevanten Fragen-Merkmal-Paare sowie zur Einschätzung der für die zu erhebungsrelevanten Merkmale erwartbaren Ausprägungen genutzt. Die gleiche Technik der paarweisen Gegenüberstellung in einer übersichtlichen Matrize wird verwandt, um die Beziehungen zwischen den zentralen Kategorien der einzelnen Forschungsschritte explizit begründet einzuschätzen - siehe die folgende Abbildung, „Verbindung der Forschungsschritte mit Hilfe von Entscheidungsmatrizen“.

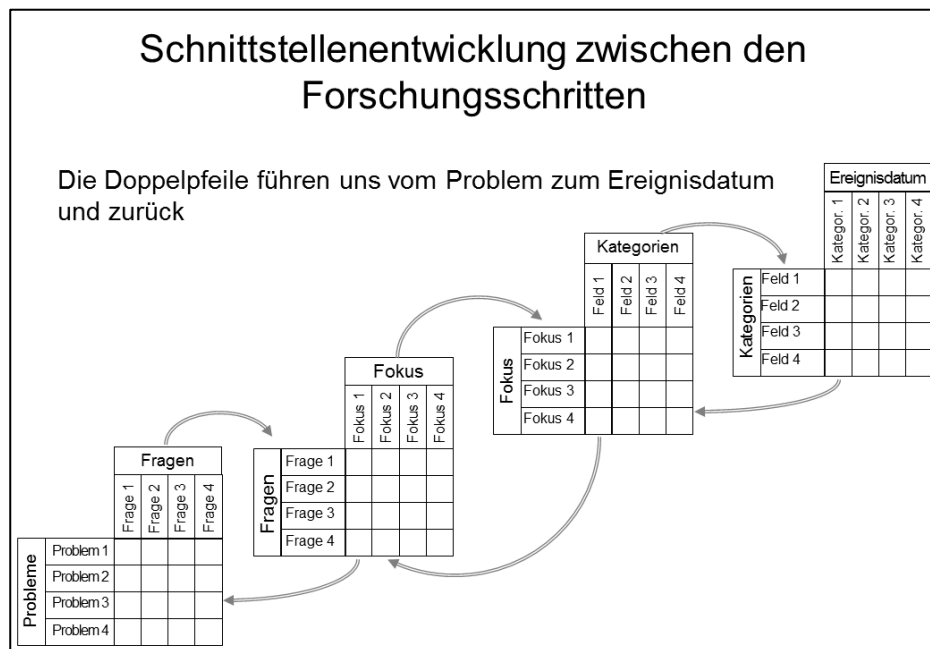


Abbildung 3: Verbindung der Forschungsschritte mit Hilfe von Entscheidungsmatrizen

Die Matrizentechnik erleichtert die Kalkulation der Beziehungen zwischen den einzelnen Forschungsschritten explizit - auch theoriegestützt und / oder in Expertenzirkeln. Dadurch lassen sich erwartungswidrige Ergebnisse rückwirkend zur Auffindung der logischen Fehlannahmen nutzen - und daraus lernen. Die Popper'sche Gegenüberstellung von „kühnen Hypothesen“ und „schärfster Kritik“ (Popper 2004, S. 118) von einfallsreichem Wirkungsmodell und empirischem Befund löst sich damit in eine *nachvollziehbare Kette begründeter Wirkungsannahmen* auf. Das erweiterte Kalkül ergänzt die direkte Ursache-Wirkungs-Beziehung um zwei Ursache-Wirkungs-Überlegungen. Vom Ergebnis her gedacht wird zuerst die zusätzliche Frage beantwortet: Was sind die notwendigen Bedingungen um das Ergebnis zu erzielen? Nun wird ein Bindeglied zwischen diesen Bedingungen und den vorhandenen Handlungsmöglich-

keiten geknüpft: Welchen Einfluss⁹ haben die Maßnahmen auf diese Bedingungs Momente?

Wirkungsannahmen kommen im Forschungsprozess eine herausgehobene Bedeutung zu. Zur Explizierung der Wirkungsannahmen empfiehlt sich die Nutzung des *Wirkungsmodell-Schemas* - siehe die folgende Abbildung. In einem Forschungsprozess begegnen sich zumindest zwei Wirkungs-Logiken, die Eigenlogik des Feldes (Forschungsgegenstand) und die Eigenlogik der Forschungsprozedur (Forschungskette). Um hierzu einen Überblick zu gewinnen, muss zuerst eine explizite Vorstellung von den *Wirkungsmomenten des zu untersuchenden Feldes* entwickelt werden. Dies kann beispielsweise mit dem Ziel der Akteure in der zu untersuchenden Situation beginnen - siehe beispielhaft die folgende Abbildung. Etwa: Welches Ziel verfolgt eine Weiterbildungsmaßnahme zur Vertiefung der pädagogischen Beobachtungskompetenz? Welche Gelingsbedingungen müssen zur Erreichung dieser Ziele gegeben sein? Dann werden die Einflussmöglichkeiten der beteiligten Akteure der ausgewählten Situation ermittelt. Zuletzt beginnt die Frage nach den von den ForscherInnen beobachtbaren Handlungen oder Maßnahmen, die die unterstellten Einflussmöglichkeiten auf die Gelingsbedingungen realisieren. Zur Erhöhung der logischen Stringenz sollte nun rückwärts von den Maßnahmen aus die Wirkungskette in Richtung Ziel untersucht werden: Wenn diese Maßnahme durchgeführt wird, realisiert sich dann notwendig der gewünschte Einfluss? Und, wenn dieser Einfluss ausgeübt wird, bilden sich dann notwendig die notwendigen Gelingsbedingungen? Wenn diese Gelingsbedingungen gegeben sind, wird dann notwendig die gewünschte Wirkung erreicht?

Die Entwicklung eines solchen Wirkungsmodells offenbart nicht nur die eigenen Überzeugungen, sondern hilft auch Wissenslücken festzustellen, die in der Wissensaufbereitung / dem „Theorieteil“ noch nicht ausreichend recherchiert wurden. Das Wirkungsmodell entwickelt sich immer weiter und steht in der empirischen Erhebung vor der seiner Prüfung. Ähnlich verfährt man mit dem *Wirkungsmodell für die Eigenlogik der Forschungskette*: Welches Problem - um im Weiterbildungsbeispiel zu bleiben - will das Forschungsvorhaben durch seine Erhebungen mit welchem Vorgehen lösen? Welchen Sinn machen diese Untersuchungen für die Feld-AkteurInnen? Auf diesen Wirkungsannahmen bauend, folgt schließlich die konkrete Eingrenzung des Problembündels auf durch die mögliche Untersuchung realistisch lösbare Aspekte.

⁹ Niklas Luhmann ist einer der wenigen Systemtheoretiker, die sich ausdrücklich mit der Kategorie „Einfluss“ in Organisationen beschäftigten. Systemisch gesehen wirkt Einfluss weniger im Sinne der linearen Handlungslogik als einseitige Einwirkung des Beeinflussenden auf das Beeinflusste (Luhmann 1995).

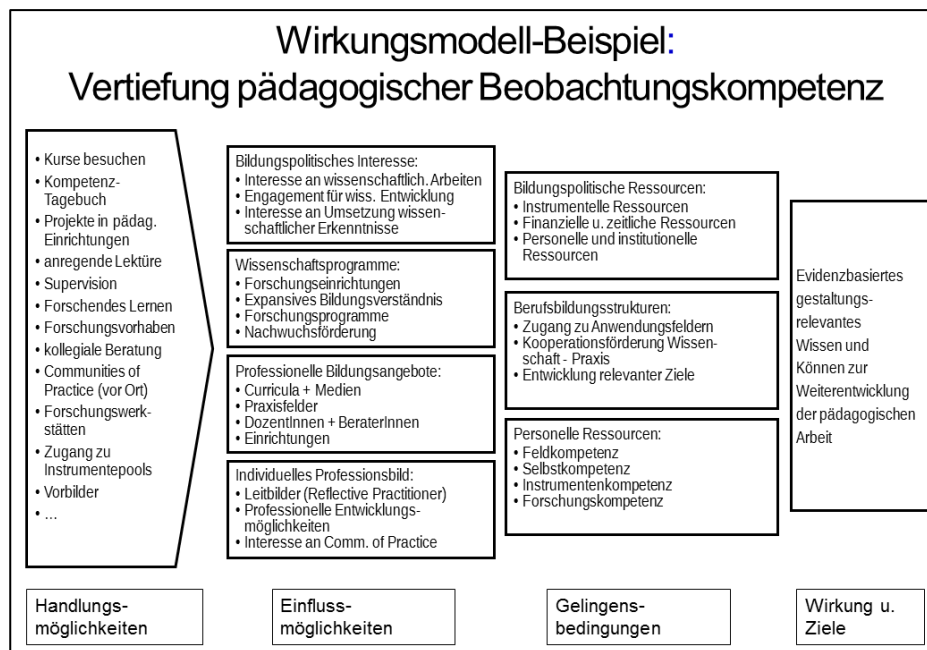


Abbildung 4: Beispiel für die Verwendung des Wirkungsmodell-Schemas

In einem durch die Hypothesenentwicklung per Entscheidungsmatrix, durch die Schnittstellenentwicklung zwischen den Forschungsschritten und durch die Wirkungsmodellierung theoretisch geklärten und sachlogisch verknüpften Forschungsprozess erübrigt sich auch die Frage nach qualitativen oder quantitativen, nach gemischten oder triangulierten Erhebungs- und Auswertungsverfahren, nach Feld- oder Laborerhebungen, der Fremd- oder Selbsterhebung usw. Die *Beantwortung der Methodenfrage* folgt der Gegenstandsangemessenheit und der Forschungsprozesslogik. Allenfalls stellt sich die Frage, ob die traditionellen Erhebungsverfahren überhaupt ausreichen oder ob nicht neue, komplexere Verfahren adaptiert oder entwickelt werden müssen - so wie Ursula Carle vor 17-18 Jahren bei ihrer Untersuchung der Schulentwicklungsmöglichkeiten verfuhr (dies., 2000, Kapitel 10, Projektperspektive: Methodeninnovation zur Sicherung schulischer Entwicklungsprojekte, S. 381 ff).

Schluss

Ausgangspunkt dieses methodologischen Entwurfs bildet das Problem, dass Fortschritte bei der sozialen und pädagogischen Inklusion auch der Unterstützung durch die empirisch-pädagogische Forschung bedürfen und diese Unterstützung aus unterschiedlichen Gründen nur unzureichend geleistet wird. Das wirkungskritische Hauptmotiv findet sich überraschenderweise in beiden Bereichen wieder, in der Integrationspädagogik wie in der empirisch-pädagogi-

sche Forschung: Die „Vorteile“ der Segregation überwiegen immer noch. Die Kosten für die Umsetzung der Integration werden zu hoch eingeschätzt, obwohl die Zusammenarbeit von WissenschaftlerInnen und ExpertInnen aus dem Praxisfeld zur Erreichung abgestimmter Entwicklungsziele Gewinn verspricht. Die Überwindung der „Methodologischen Segregation“ ist in der pädagogischen Forschung wohl immer noch „zu teuer“, zu aufwendig, zu fremd. Dass und wie sie dennoch methodologisch prinzipiell möglich ist, versuchte dieser Artikel entsprechend den vier Kernbereichen eines methodologischen Paradigmas zu beschreiben: (1) die notwendigen wissenschaftsphilosophischen Annahmen und Haltungen, (2) die neue Untersuchungslogik (Methodologie), (3) einige Methoden, Techniken und Leitlinien für die Forschungspraxis. Der Artikelökonomie geschuldet, fehlt noch der vierte Bereich, (4) die sozialen und politischen Verpflichtungen zu Forschung in sozialer und gesellschaftlicher Verantwortung. So bleibt einiges offen, was anregend ist für die Diskussion und motivierend für detaillierte und tiefer gehende Ausführungen. Etwa die Frage danach, wer die eingeschränkte Lernqualität der szientistischen wie der schulisch-unterrichtlichen Segregation eigentlich noch will und wer nicht mehr.

Literatur

- Arce-Gomez, Antonio; Donovan, Jerome D.; Bedggood, Rowan E. (2015): Social impact assessments. Developing a consolidated conceptual framework. In: *Environmental Impact Assessment Review*, 50, S. 85–94. Online verfügbar unter <http://www.sciencedirect.com/science/article/pii/S0195925514000742>, zuletzt geprüft am 07.09.2016.
- Campbell, Donald T. (1988): *Methodology and epistemology for social science. Selected papers.* Edited by E. Samuel Overman. Chicago: University of Chicago Press.
- Carle, Ursula (o. J.): *Arbeitsgebiet Elementar- und Grundschulpädagogik.* Webseite der Universität Bremen. Themenseite "Forschung" Bremen: Universität Bremen. Online verfügbar unter <http://www.grundschulpaedagogik.uni-bremen.de/forschung/>, zuletzt geprüft am 07.09.2016.
- Carle, Ursula; Berthold, Barbara (2004): *Schuleingangsphase entwickeln – Leistung fördern. Wie 15 Staatliche Grundschulen in Thüringen die flexible, jahrgangsgemischte und integrative Schuleingangsphase einrichten.* Baltmannsweiler: Schneider Hohengehren
- Carle, Ursula; Samuel, Annette (2007): *Frühes Lernen - Kindergarten und Grundschule kooperieren.* Baltmannsweiler: Schneider Hohengehren
- Carle, Ursula; Metzen, Heinz (2014): *Wie wirkt Jahrgangsübergreifendes Lernen? Internationale Literaturübersicht zum Stand der Forschung, der praktischen Expertise und der pädagogischen Theorie. Eine wissenschaftliche Expertise des Grundschulverbandes.* Frankfurt am Main: Grundschulverband (GSV) (Wissenschaftliche Expertisen). Online verfügbar unter <http://www.grundschulverband.de/veroeffentlichungen/wissenschaftliche-expertisen/>, zuletzt geprüft am 20160716.

-
- Carle, Ursula; Metzen, Heinz (2013): Evaluation und Frühpädagogik: Entwicklung der Evaluationsmethodik für frühkindliche Bildungsprogramme. In Margit Stamm; Doris Edelmann (Hrsg.), *Handbuch frühkindliche Bildungsforschung* (S. 857–873). Wiesbaden: Springer VS.
- Carle, Ursula; Metzen, Heinz (2009): Die Schuleingangsphase lohnt sich! Erfolgsmomente für die bestmögliche Entwicklung des landesweiten Schulentwicklungsvorhabens 'Begleitete Schuleingangsphase' in Thüringen. Bericht der wissenschaftlichen Begleitung nach zweieinhalb Jahren 'BeSTe' (2005-2008). Bremen: Universität Bremen (Internetpublikation) URL: [http://www.grundschulpaedagogik.uni-bremen.de/schuleingangsphase/2009_02beste_endbericht\(gesamt\).pdf](http://www.grundschulpaedagogik.uni-bremen.de/schuleingangsphase/2009_02beste_endbericht(gesamt).pdf). Download 20160912
- Carle, Ursula; Metzen, Heinz (2006): So lassen sich Netzwerke (nicht) evaluieren. Von der Netzwerkrhetorik zur systemischen Organisationsgestaltung und Forschungsmethodik - erste Gehversuche bei Vernetzungsversuchen in der Familienbildung. Manuskript zum Vortrag, gehalten im Workshop des AK Soziale Dienstleistungen der DeGEval, "Evaluation von Netzwerken in der sozialen Arbeit: Probleme, methodische Herausforderungen und konzeptionelle Antworten (A2)", Lüneburg. Online verfügbar unter http://shopfloor.de/evaluation/lueneburg_artikel.html, zuletzt aktualisiert am 2006, zuletzt geprüft am 16.07.2016.
- Collins, Allan; Joseph, Diana; Bielaczyc, Katerine (2004): Design Research. Theoretical and Methodological Issues. In: *Journal of the Learning Sciences* 13 (1), S. 15–42.
- Della-Piana, Gabriel; Kubo Della-Piana, Connie; Gardner, Michael K. (2015): Sources of Influence on the Problem of a Validity Evidence Gap for Education Achievement Tests. In: *Teachers College Record (TCR)* (ID Number: 18064).
- Dewey, John (2002, 1938): *Logik. Die Theorie der Forschung*. Aus dem Englischen von Martin Stühr. *Englisches Original, zuerst 1938: Logic. The Theory of Inquiry*. Carbondale: Southern Illinois University Press). Frankfurt am Main: Suhrkamp.
- Feyerabend, Paul K. (2002, 1965): *Probleme des Empirismus*. Teil 1. Aus dem Englischen übersetzt und mit einem Nachwort von Volker Böhnigk und Rainer Noske. Stuttgart: Reclam (Universal-Bibliothek, 18139). *Englisches Original, zuerst 1965, zweite Auflage 1983: Problems of Empiricism*. In: Robert Garland Colodny (Ed.): *Beyond the edge of certainty. Essays in contemporary science and philosophy*. Lanham, MD, USA: University Press of America (CPS publications in philosophy of science), S. 145–260.
- Foerster, Heinz von; (1993): *Kybernetik*. Autorisierte Übersetzung aus dem Amerikanischen von Birger Ollrogge. Berlin: Merve-Verlag (Internationaler Merve-Diskurs & Perspektiven der Technokultur, 180).
- Friedrichs, Jürgen (1999): *Methoden empirischer Sozialforschung*. 15. Aufl. Opladen: Westdeutscher Verlag (WV-Studium Sozialwissenschaft, 28).
- Green, Jennifer C. (2006): Toward a methodology of mixed methods social inquiry. In: *Research in the Schools (RITS)* 13 (1), S. 93–98. Online verfügbar unter <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.462.9391&rep=rep1&type=pdf>, zuletzt geprüft am 11.09.2016.

- Grossmann, Claudia; Sanders, Julia; English, Rebecca A. (2013): Large Simple Trials and Knowledge Generation in a Learning Health System. Workshop Summary. Washington, DC, USA: National Academies Press. Online verfügbar unter http://www.nap.edu/catalog.php?record_id=18400, zuletzt geprüft am 10.09.2016.
- Hattie, John (2013): Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von Visible Learning besorgt von Wolfgang Beywl und Klaus Zierer. Baltmannsweiler: Schneider Verlag Hohengehren.
- Hoyos, Carl Graf; Metzen, Heinz (1982): Belastung und Beanspruchung bei Steuerungs- und Überwachungstätigkeiten. Untersuchungen in Fahrdienstleitungen der Deutschen Bundesbahn. Unter Mitarbeit von H. Dvorak, E. Hausmann und R. Leitmeier (Berichte aus dem Institut für Psychologie und Erziehungswissenschaft der Technischen Universität München, 9). München: Institut für Psychologie und Erziehungswissenschaft der Technischen Universität München.
- Luhmann, Niklas (1995-1964): Funktion und Folgen formaler Organisation. Mit einem Epilog 1994 (4. Aufl.). Berlin: Duncker & Humblot.
- Metzen, Heinz (1977): Empirische Sozialforschung in der kulturellen Außenpolitik der Bundesrepublik Deutschland. Möglichkeiten und Maßstäbe. Zugl. Dissertation, Universität München. München: Selbstverlag.
- Nida-Rümelin, Julian (2011): Verantwortung. Stuttgart: Reclam (Universal-Bibliothek, 18829).
- Onwuegbuzie, Anthony J.; Johnson, R. Burke; Collins, Kathleen M. T. (2009): Assessing legitimacy in mixed research. A new framework. In: *Quality & Quantity (Q&Q)* 45 (6), S. 1253–1271.
- Patton, Michael Q. (2010). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York: Guilford Press.
- Pawson, Ray (2013): *The science of evaluation. A realist manifesto*. Los Angeles, CA: Sage.
- Pawson, Ray; Tilley, Nick (2011, 1997): *Realistic evaluation*. Reprinted. London: Sage.
- Pawson, Ray (2003): Nothing as Practical as a Good Theory. This contribution is based on the second plenary address given at the 5th biennial meeting of the European Evaluation Society, 12 October 2002 in Seville, Spain. In: *Evaluation* 9 (4), S. 471–490. Online verfügbar unter <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.116.7540&rep=rep1&type=pdf>, zuletzt geprüft am 16.09.2016.
- Peter, Sophie (2016): Understanding and assessing the societal impact of sustainability research: From state of the art to a process-oriented assessment plan. With the case study of the Wuppertal Institute in Germany. Thesis for the fulfilment of the Master of Science in Environmental Sciences, Policy & Management (MESPOM), Lund University. Lund, SWE: Lund University Libraries (IIIEE Masters Thesis, IMEN56 20161). Online verfügbar unter <http://lup.lub.lu.se/student-papers/record/8890692>, zuletzt geprüft am 07.09.2016.

-
- Popper, Karl R. (2004): Ausgangspunkte. Meine intellektuelle Entwicklung. Unter Mitarbeit von Friedrich Griese. Ungekürzte Taschenbuchausgabe. München u.a.: Piper (Serie Piper, 4188).
- Scriven, Michael (2004). The Fiefdom Problem: Editorial. *Journal of MultiDisciplinary Evaluation (JMDE)*, 1(1), 11–18. Online verfügbar unter http://journals.sfu.ca/jmde/index.php/jmde_1/issue/view/19, zuletzt geprüft am 11.09.2016.
- Stamm, Margit; Edelmann, Doris (Hrsg.) (2013): *Handbuch frühkindliche Bildungsforschung*. Wiesbaden: Springer VS.
- Sylva, Kathy; Melhuish, Edward; Sammons, Pam; Siraj, Iram; Taggart, Brenda (2014): Students' educational and developmental outcomes at age 16. Effective pre-school, primary and secondary education (EPPSE 3-16) project. Unter Mitarbeit von Rebecca Smees, Katalin Toth, Wesley Welcomme und Katie Hollingworth. London: Institute of Education, University of London, GBR (Research Report, 354). Online verfügbar unter <http://www.ucl.ac.uk/ioe/research/pdf/16-educational-Developmental-Outcomes-RR.pdf>, zuletzt geprüft am 07.05.2016.
- Thurston, Wilfreda E.; Potvin, Louise (2003): Evaluability Assessment. A Tool for Incorporating Evaluation in Social Change Programmes. *Evaluation*, 9 S. (4), 453–469. Online verfügbar unter <http://evi.sagepub.com/content/9/4/453>, zuletzt geprüft am 09.09.2016.
- Villar, José; Duley, Lelia (2003): The need for large and simple randomized trials in reproductive health. UNDP / UNFPA / WHO / World Bank Special Programme of Research, Development and Research Training in Human Reproduction, WHO, Geneva, Switzerland and Resource Centre for Randomised Trials, Institute of Health Sciences Headington, Oxford, United Kingdom. In: *The WHO Reproductive Health Library*, No 6. Online verfügbar unter http://apps.who.int/rhl/The_need_for_large_and_simple_randomized_trials.pdf, zuletzt geprüft am 10.09.2016.