



Dirk Nüsken | Wolfgang Böttcher

Was leisten die Erziehungshilfen?

Eine einführende Übersicht zu
Studien und Evaluationen der HzE

Dirk Nüsken | Wolfgang Böttcher
Was leisten die Erziehungshilfen?

Basistexte Erziehungshilfen

Herausgegeben im Auftrag der
Internationalen Gesellschaft für erzieherische
Hilfen (IGfH) von

Josef Koch | Friedhelm Peters |

Elke Steinbacher | Wolfgang Trede

Dirk Nüsken | Wolfgang Böttcher

Was leisten die Erziehungshilfen?

Eine einführende Übersicht
zu Studien und Evaluationen der HzE

Die Autoren

Dirk Nüsken, Jg. 1970, Dr. phil., Diplomsozialpädagoge, ist Professor für Theorie und Praxis der Sozialen Arbeit mit dem Schwerpunkt Kinder- und Jugendhilfe an der Evangelischen Hochschule Rheinland-Westfalen-Lippe in Bochum.

Wolfgang Böttcher, Jg. 1953, Dipl. rer. soc., Dr. rer. pol. habil., ist Professor für Erziehungswissenschaft mit den Schwerpunkten Qualitätsentwicklung und Evaluation in Einrichtungen des Bildungs- und Sozialwesens an der Westfälischen Wilhelms-Universität (WWU) Münster.

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung ist ohne Zustimmung des Verlags unzulässig. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronische Systeme.



Dieses Buch ist erhältlich als:
ISBN 978-3-7799-2693-1 Print
ISBN 978-3-7799-5197-1 E-Book (PDF)

1. Auflage 2018

© 2018 Beltz Juventa
in der Verlagsgruppe Beltz · Weinheim Basel
Werderstraße 10, 69469 Weinheim
Alle Rechte vorbehalten

Herstellung: Ulrike Poppel
Satz: text plus form, Dresden
Druck und Bindung: Beltz Grafische Betriebe, Bad Langensalza
Printed in Germany

Weitere Informationen zu unseren Autor_innen und Titeln finden Sie unter:
www.beltz.de

Zu diesem Buch

Die Frage nach den Ergebnissen oder Wirkungen pädagogischer oder sozialpädagogischer Maßnahmen bestimmt aktuell zahlreiche fachliche und fachpolitische Diskurse. Sie wird nicht nur – aber eben auch – an die Hilfen zur Erziehung (HzE) gerichtet: Was bewirkt der Einsatz von Geld und Motivation, Arbeitszeit und gutem Willen in den Hilfen für Kinder, Jugendliche und Eltern tatsächlich? Angesichts der Dominanz von ökonomischen Fragestellungen, die im Rahmen des New Public Management auch an Leistungen der Sozialen Arbeit gestellt werden, gilt es, diese Frage fachlich angemessen und so weitreichend wie möglich zu beantworten. Dazu bedarf es entsprechender empirischer und bewertender Forschung.

Dem entsprechend wollen wir eine grundlegende Einführung in Studien und Evaluationen der Hilfen zur Erziehung geben. Interessierte Leser_innen erhalten hier einen systematischen Überblick zu empirischen Arbeiten aus etwa sechs Jahrzehnten, die zeigen, dass inzwischen durchaus ein gewisser Fundus an empirischen Erkenntnissen vorliegt.

Im Mittelpunkt stehen ausgewählte Studien ab den 1950er Jahren, die sich mit den HzE und hier insbesondere mit den Erfolgen von Hilfemaßnahmen und den Entwicklungsprozessen junger Menschen innerhalb bzw. nach Abschluss erzieherischer Hilfen beschäftigen. Mit der vorgenommenen Auswahl wollen wir einerseits grundsätzliche Entwicklungen (Studiendesigns, Bewertungskriterien etc.) aufzeigen, zum anderen sollen Kontraste bestimmter Zugänge deutlich werden, so dass Unterscheidungen und Bewertungen solcher Erhebungen und ihrer Ergebnisse möglich sind. Wir referieren dazu die Grundfragen, die die jeweiligen Studien stellen, die Perspektiven und Quellen der erhobenen Daten sowie die jeweiligen Kriterien, die in den Studien zur Operationalisierung der Fragestellungen angelegt werden. Neben den unterschiedlichen methodologischen und methodischen Zugän-

gen werden dabei auch die durchaus verschiedenen Funktionen, die den HzE in den jeweiligen Studien zugeschrieben werden, deutlich. Im Mittelpunkt der Mehrheit der Studien stehen Maßnahmen der Heimerziehung. Um die Breite des Leistungsfeldes zu verdeutlichen, zeigen wir jedoch auch einige Beispiele auf, in denen es um die Analyse von teilstationären und ambulanten Erziehungshilfen und um Pflegefamilien geht. Neben der Beschreibung und kritischen Würdigung jeder Studie stellen wir zugleich heraus, welche Empfehlungen zur Qualitätsentwicklung und Professionalisierung der HzE aus den jeweiligen Studienergebnissen abgeleitet werden bzw. werden können. Knappe Zwischenfazits zu bestimmten Studientypen sowie Hinweise auf vorliegende Übersichten, Zusammenfassungen und Metaanalysen fließen an den entsprechenden Stellen mit ein und bieten, gemeinsam mit ausgewählten Literaturhinweisen, Möglichkeiten zur vertiefenden Auseinandersetzung. Eine kurze Zusammenfassung jeder dargestellten Studie findet sich jeweils am Ende unter der Überschrift „Auf einen Blick“.

Im abschließenden Kapitel bilanzieren wir unsere Befunde und formulieren gebündelte Schlussfolgerungen aus den analysierten Studien. Wir fragen dabei auch nach der Bedeutung von Ergebnissen aus Studien und Evaluationen für die Praxis und Politik der HzE. Ein weiterer Blick gilt den Funktionen und den Prinzipien guter Evaluationen.

Das Buch richtet sich an Fachkräfte der Hilfen zur Erziehung und an Entscheidungsträger in zuständigen Organisationen und fachpolitischen Gremien. Es soll auch dazu dienen, Studierende an Hochschulen, Akademien und Fachschulen in das Themenfeld erzieherischer Hilfen und ihrer wissenschaftlichen Analyse und Bewertung einzuführen. Wir danken Sofia Helena Peters und Daniela Engelbracht für die Unterstützung bei der formalen Gestaltung dieses Bandes.

Bochum, im Frühjahr 2018
Dirk Nüsken, Wolfgang Böttcher

Inhalt

1.	Zur Einführung: Hilfen zur Erziehung und deren Bewertung	9
2.	Studien der letzten Jahrzehnte und methodische Zugänge	24
2.1	Lebensbewährungsstudien – Ein historischer Zugang	27
2.2	Studien zur Lebenssituation von Heim- und Pflegekindern	36
2.3	Hilfverlaufsstudien	48
2.4	Evaluative Controlling- und Qualitätsentwicklungsinstrumente	90
2.5	Kosten-Nutzen-Studien	103
	Zwischenfazit I: Ergebnisse quantitativer Studien und Instrumente	108
2.6	Wissenschaftliche Bewertung von Wirkungen	112
	Zwischenfazit II: Wirkungsevaluationen	131
2.8	Fallstudien	138
	Zwischenfazit III: Fallstudien und Evaluation	164
2.8	Katamnestische Adressat_innenbefragungen	166
	Zwischenfazit IV: Zur Bedeutung der Adressat_innenperspektive	178
3.	Studien zu HzE – Bilanzen und Ausblicke	182
3.1	Übergreifende Einsichten und Schlussfolgerungen	184
3.2	Forschung und Partei – Forschung als Partei	192
3.3	Evaluation als Forschungsprinzip	195
3.4	Zur Bedeutung von wissenschaftlich generiertem Wissen für die sozialpädagogische Praxis und Politik	202

Literatur	207
Anlage 1	225
Verzeichnis aller in diesem Band dargestellten und kommentierten Studien und Qualitätsdokumentationsinstrumente	225

1. Zur Einführung: Hilfen zur Erziehung und deren Bewertung

Als „Hilfen zur Erziehung“ (HzE) werden seit Einführung des SGB VIII (Kinder- und Jugendhilfegesetz) 1990/1991 Beratungs-, Betreuungs- und Hilfearrangements für junge Menschen und ihre Familien verstanden, wie sie in den §§ 27 bis 35 des SGB VIII vorgesehen sind. Diese Hilfen gelten Familien mit erzieherischen Problemen, die diese nicht alleine oder mit Hilfe etwa der Kindertagesstätte oder der Schule lösen können. Leistungsträger ist der soziale Dienst des zuständigen Stadt- oder Kreisjugendamtes. Konkret geleistet werden erzieherische Hilfen jedoch zumeist von freien Trägern der Kinder- und Jugendhilfe. Hilfen zur Erziehung können in ambulanter (z. B. Sozialpädagogische Familienhilfe), teilstationärer (z. B. Tagesgruppen) oder stationärer Form (z. B. Heimerziehung oder Pflegefamilien) wie auch in integrierten Formen geleistet werden. Sie übernehmen damit familienunterstützende, familienergänzende oder familienersetzende Funktion, wobei die Ursprungsfamilie auch bei stationären Hilfen Einbindung findet und somit als solche natürlich nicht ersetzt wird und auch nicht ersetzt werden soll. Erzieherische Hilfen können von den Sorgeberechtigten freiwillig beantragt, von den Sozialdiensten „dringend empfohlen“ (und ebenfalls „freiwillig“ beantragt) oder von den Familiengerichten veranlasst werden. In den Hilfen zur Erziehung bündeln sich so zahlreiche und z. T. widersprüchliche Facetten: Erzieherische Hilfen können für Kinder oder Jugendliche Ruhe und Unterstützung, Schutz und Hilfe, Bildungsorte und neue Chancen, in nicht wenigen Fällen sogar Rettung aus unerträglichen Verhältnissen wie Vernachlässigung und Misshandlung bedeuten. Zugleich stellen sie für die betroffenen jungen Menschen und deren Eltern subjektiv betrachtet oftmals eine Bedrohung dar, d. h. sie werden als Eingriff, Kontrolle und Disziplinierung verstanden (vgl. Trede 2009, S. 15).

Gesellschaftliche Erwartungen an die Jugendämter und die Hilfen zur Erziehung werden besonders dann deutlich, wenn die Akteure des Hilfe- und Schutzsystems an ihre Grenzen gelangen und zuweilen auch versagen. Das haben die Diskurse und die medialen Reaktionen im Fall der im Kontext der HzE zu Tode gekommenen Kinder in den letzten Jahren ebenso gezeigt, wie die Ergebnisse der ‚Runden Tische zur Geschichte der Heimerziehung und zum sexuellen Missbrauch‘ (AGJ 2010; BMJ/BMF/SFJ/BMBF 2011) sowie die jüngeren Diskussionen um Skandale in der Heimerziehung (Haasenburg/Friesenhof u. a.). Deshalb sind HzE oftmals auch im Kontext der Bestimmung von Kindeswohl und Kindeswohlgefährdung zu betrachten und damit innerhalb der Balancen von Elternrecht, Kindeswohl und staatlichem Schutzauftrag.

Basisinfo: HzE und Kindeswohlgefährdung

Der Begriff der Kindeswohlgefährdung entstammt dem Kindschaftsrecht des Bürgerlichen Gesetzbuches (BGB) und knüpft an den staatlichen Schutzauftrag nach Art. 6 Abs. 2 Grundgesetz (GG) an. § 1666 Absatz 1 BGB führt als Gefährdungstatsachen die missbräuchliche Ausübung der elterlichen Sorge, die Vernachlässigung des Kindes, das unverschuldete Elternversagen und das Verhalten einer/eines Dritten auf. Sozialwissenschaftlich betrachtet lässt sich jedoch eher von Formen der Kindwohlgefährdung, insbesondere von Vernachlässigung, Misshandlung und sexueller Gewalt sprechen. Auch wenn die Bestimmung einer Kindeswohlgefährdung im Einzelfall stets ein soziales und rechtliches Konstrukt bleibt, so hat die Rechtsprechung des Bundesgerichtshofs den Begriff der Kindeswohlgefährdung insoweit konkretisiert, als sie darunter folgendes versteht: „eine gegenwärtige, in einem solchen Maße vorhandene Gefahr, dass sich bei der weiteren Entwicklung eine erhebliche Schädigung mit ziemlicher Sicherheit voraussehen lässt“ (BGH FamRZ 1956, S. 350).

Für die Fachkräfte des Sozialdienstes eines Jugendamtes ist der Rechtsbegriff „Kindeswohlgefährdung“ Maßgabe, um über Hilfsangebote und ggf. die Einschaltung des Familiengerichtes zu entscheiden. Ent-

sprechende Prozesse normiert seit 2005 § 8a SGB VIII. Dabei kommt insbesondere der Bereitschaft und Fähigkeit der Eltern eine besondere Bedeutung zu, eine Gefährdung abzuwenden und ggf. eine Hilfe anzunehmen. Von der Kindeswohlgefährdung zu unterscheiden ist jedoch ein erzieherischer Bedarf, der gem. § 27 SGB VIII gegeben ist, wenn „eine dem Wohl eines Kindes oder eines Jugendlichen entsprechende Erziehung nicht gewährleistet ist. In diesem Fall besteht ein Anspruch auf Hilfen zur Erziehung nach §§ 27 bis 35 SGB VIII. Mit der Nichtgewährleistung einer dem Kindeswohl entsprechenden Erziehung normiert das SGB VIII seit seinem Inkrafttreten 1990/1991 eine Schwelle unterhalb der Kindeswohlgefährdung nach § 8a Abs. 1 Satz 1 SGB VIII, § 1666 Abs. 1 BGB“ (Schmid/Meysen 2007). Eine Intervention des Jugendamtes an dieser Schwelle setzt das Einverständnis der Eltern voraus. Erst bei Vorliegen einer Kindeswohlgefährdung und wenn Eltern nicht bereit und in der Lage sind, eine Gefährdung abzuwenden und z. B. erzieherische Hilfen in Anspruch zu nehmen, muss das Jugendamt das Familiengericht anrufen, welches Maßnahmen gegenüber den Eltern (Gebote/ Verbote) ergreifen und auch in das Sorgerecht eingreifen kann (vgl. ebd.).

Das Interesse an der Bewertung von Hilfen zur Erziehung und insbesondere der Blick auf Ergebnisse stehen bereits seit langem im Mittelpunkt zahlreicher Studien.

Vor allem die Frage danach, was als Ergebnis oder Erfolg verstanden werden kann, welche Kriterien und Erhebungsmethoden der Erziehungshilfe angemessen sind und wozu solch einschlägiges Wissen dienen kann, wird durchaus unterschiedlich diskutiert. Dabei geht es um die Bewertung von HzE. So zeigt etwa der in den letzten Jahren geführte Diskurs um die Frage nach den Wirkungen pädagogischer oder sozialpädagogischer Maßnahmen im Allgemeinen (vgl. Kapitel 2.6) und um eine Wirkungsorientierung in der Erziehungshilfe im Besonderen die Unterschiedlichkeit von Grundannahmen auf. Mit Blick auf eine an den Ergebnissen und Erkenntnissen empirischer Forschung ausgerichtete Erziehungshilfe kritisiert Peters den Ansatz einer Sozialtechnologie, also den Versuch, durch die Bestimmung von Ursache-Wirkungszusammenhängen möglichst exakte Bestimmungen von

Mitteln und Wegen zu bestimmten, vorab definierten Zielen zu machen. Mit Bezug auf Thiersch (1993, S. 179) verweist Peters darauf, dass solche kausalen Zusammenhänge im Feld einer lebensweltorientierten Sozialen Arbeit nicht herstellbar seien, da Sozialpädagogik zwar anregen und unterstützen könne, nicht aber dafür verantwortlich sei, was die Adressat_innen daraus machen (Peters 2006, S. 207 ff.). Der Verein für Kommunalwissenschaft fragt gar, ob wirkungsorientierte Steuerung nicht lediglich ein Mythos – eine Sage, eine Vorstellung, der man keinen Glauben schenken darf – ist (VfK 2008)? Die Notwendigkeit wie die Selbstverständlichkeit der Messbarkeit der Wirkungen von HzE betonen dagegen Lambach (2003), Macsenaere (2006) und Schröder (2002). Tornow (2005) wie auch Kleinrahm et al. (2012) befürworten zudem eine wirkungsorientierte Steuerung der Leistungen im Rahmen der HzE. Welche Bedeutung empirisch gewonnene Erkenntnisse für die Entwicklung von Konzepten und für das alltägliche Handeln von Fachkräften und damit für eine Professionalisierung haben können, thematisieren Albus et al. (2010), Nüsken (2011a) und Schäfer/Pierlings (2011).

Betrachtet man die Grundidee hinter der Frage nach Ergebnissen oder Erfolgen der HzE, so zeigt sich, dass es beim sozialpädagogischen Handeln darum geht, Veränderungen bzw. Wirkungen im Leben von Kindern, Jugendlichen und ihren Familien zu erzielen. Dementsprechend ist es legitim, danach zu fragen, ob oder inwieweit anvisierte Ziele auch erreicht werden. Und da praktische sozialpädagogische Arbeit auch den Einsatz von Ressourcen – Geld, Zeit, Wissen oder Engagement – erfordert, ist auch die Frage verständlich, ob das Verhältnis von Aufwand und Ergebnis angemessen ist. In diesem Kontext und ausgehend von den USA und dem UK, unterstützt durch die OECD als „Global Player“ für eine „ökonomische“ Ausrichtung von Bildung und Sozialer Arbeit, hat die Frage nach den Ergebnissen und Effekten von Leistungen im Sozial- und Bildungswesen eine erfolgreiche Reise um die Welt absolviert (vgl. Böttcher/Nüsken 2015, S. 349).

Nicht unterschlagen werden darf mit Blick auf den deutschen Diskurs, dass das verstärkte Interesse an den Wirkungen erzieherischer Hilfen in den letzten Jahren auch im Zusammenhang mit

dem Ausgabenanstieg der öffentlichen Jugendhilfe für Maßnahmen der HzE und der problematischen Lage der öffentlichen Haushalte zu sehen ist (Struzyna 2007).¹ Insofern erstaunt es nicht, dass sich eine Reihe von Studien (zumindest auch) als strategische Fortführung des Qualitätsdiskurses in der Jugendhilfe lesen lassen (vgl. Haubrich 2009, S. 445). Dem „klassischen“ Dreischnitt von Struktur-, Prozess- und Ergebnisqualität (Donabedian 1980) folgend, geraten Aspekte der Ergebnisse dabei in den Mittelpunkt. Dementsprechend lässt sich die Zunahme von Studien in sozialpädagogischen Leistungsfeldern in Deutschland zunächst vor allem als eine Weiterentwicklung der Neuen Steuerungsmodelle im Kontext des New Public Management lesen, die mit dem Ziel einer Effektivitätssteigerung öffentlicher Leistungen angetreten sind (vgl. Albus et al. 2010, S. 20).

Basisinfo: New Public Management (NPM)

Das NPM ist ein Sammelbegriff für verschiedene, durch die „Vermarktlichung des Sozialstaates“ (Nullmeier 2004, S. 495) inspirierte Aktivitäten, die die Verwaltung öffentlicher Dienstleistungen modernisieren wollen. Das Modell der „bürokratischen Steuerung“, welches das „alte“ Management bestimmte, trennte die Verantwortung für die Finanzen von der eigentlichen Dienstleistung, in der Regel fand eine auf das Kalenderjahr bezogene Kontrolle und Abrechnung der Ausgaben statt, die Abwicklung der Aufgaben erfolgte im Sinne eines reibungslosen Verwaltungsablaufs. Bürokratie geht – mehr oder weniger eindeutig – von der Annahme aus, dass Handlungsanweisungen von höheren Ebenen der Verwaltungshierarchie problemlos und unverändert umgesetzt werden, die Arbeitsaufträge werden in einem hoch arbeitsteiligen Verfahren erledigt, Details werden akribisch vorgeschrieben.

1 Eine Kurzvorstellung von etwa 40 Studien zur Wirkungen bzw. zu Ergebnisbewertungen der Erziehungshilfe und zu deren z. T. sehr unterschiedlichen Zugängen findet sich in dem Band „Outcome in der Jugendhilfe gemessen“ von Macsenaere/Hiller/Fischer (2010).

Das neue Management lässt sich durch „Steuerung auf Abstand“ kennzeichnen. Es geht um die Formulierung von Zielen, die von den Handlungseinheiten weitgehend selbstständig umgesetzt werden sollen. Konkurrenz und andere Prinzipien aus dem Management profitorientierter Unternehmen wurden eingeführt, fragmentierte Aufgaben werden wieder im Sinne der „Kunden“ zusammengeführt. Die Dienste erbringenden Akteure haben erweiterte Verfügungsrechte über Budgets. Auf der anderen Seite wurden Rechenschaftssysteme eingeführt, die die Leistungsüberprüfung, Qualitätssicherung und -entwicklung zum Ziel hatten (ausführlich: Blanke et al. 2011).

Zum Tragen kam das NPM in Deutschland vor allem durch das auf kommunaler Ebene umgesetzte Neue Steuerungsmodell (NSM). Kennzeichen dieses durch die Kommunale Gemeinschaftsstelle für Verwaltungsvereinfachung (KGST) in den 1990er Jahren beförderte Modell sind die Bestimmung von Produkten und deren Zielen und Kennziffern der Erfolgskontrolle, eine dezentrale Ressourcenverwaltung sowie Kontrakte, in denen Zielvorgaben und Vergütungsregelungen festgelegt werden. Diese sind ferner einem entsprechenden Controlling unterworfen. Die Elemente des NPM bzw. NSM lassen sich insbesondere in den Regelungen zur Qualitätsentwicklung der HzE finden.

Basisinfo: Qualitätsentwicklung erzieherischer Hilfen

Am 01.01.1999 trat mit den §§ 78a–g SGB VIII eine Neuordnung der Leistungs- und Finanzierungsregelungen der Erziehungshilfen in Kraft. Mit den hier verfassten prospektiv, also für einen zukünftigen Zeitraum gedachten Leistungs-, Entgelt- und Qualitätsentwicklungsvereinbarungen zwischen öffentlichem und freiem Träger hat der Gesetzgeber drei Ziele verbunden:

- die Dämpfung der Kostenentwicklung insbesondere im teilstationären und stationären Bereich der Kinder- und Jugendhilfe,
- die Schaffung einer stärkeren Transparenz von Kosten und Leistungen,

- sowie die Verbesserung der Effizienz der eingesetzten Mittel (vgl. Bundestagsdrucksache 13/10330 v. 01.04.1998, S. 16)

Die in § 78b SGB VIII geregelten Qualitätsentwicklungsvereinbarungen sollen dazu dienen, Vereinbarungen über die „Grundsätze und Maßstäbe für die Bewertung der Qualität der Leistungsangebote sowie über geeignete Maßnahmen zu ihrer Gewährleistung“ (§ 78b SGB VIII Abs. 1, 3) abzuschließen. Mit den, im Kontext des Bundeskinderschutzgesetzes 2012, eingeführten Regelungen des § 79a SGB VIII (Qualitätsentwicklung in der Kinder- und Jugendhilfe) wurde den öffentlichen Trägern ferner aufgegeben, zur Erfüllung der Aufgaben der Jugendhilfe Maßstäbe für die Bewertung der Qualität sowie geeignete Maßnahmen zu ihrer Gewährleistung weiterzuentwickeln, anzuwenden und regelmäßig zu überprüfen. Solche Maßstäbe und Maßnahmen sollen der Gewährung und Erbringung von Leistungen, der Erfüllung anderer Aufgaben sowie dem Prozess der Gefährdungseinschätzung nach § 8a SGB VIII und der Zusammenarbeit mit anderen Institutionen gelten. Dazu zählen auch Qualitätsmerkmale für den Schutz vor Gewalt und die Sicherung der Rechte von Kindern und Jugendlichen in Einrichtungen.

Bei der Differenzierung des hier relevanten Qualitätsbegriffes hat sich die von dem Mediziner und Public Health Forscher Avedis Donabedian aufgegriffene Unterscheidung in Struktur-, Prozess- und Ergebnisqualität als zentrale Qualitätsdimensionen (auch) in der Kinder- und Jugendhilfe durchgesetzt.

Zur Strukturqualität zählen in diesem Verständnis die organisatorischen Rahmenbedingungen und die Ausstattung einer Einrichtung, wie etwa das Personal und dessen Qualifikation, die Konzepte und die Gebäude und deren Ausstattung.

Die Prozessqualität bezieht sich hingegen auf die Art und Weise einer Leistungserbringung. In den Blick genommen werden hier z. B. Zieltransparenz, Dienstplangestaltung, Teambesprechungen, Beteiligung und Beschwerdemöglichkeiten von Kindern und Jugendlichen und die Gestaltung von Interaktionen.

Im Vordergrund der Ergebnisqualität stehen die Themen Zielerreichung und Wirkung einer Maßnahme. Zeitpunkte, Kriterien und Indikatoren, die Perspektiven (Adressat_innen, Fachkräfte, externe Forscher_in-

nen) wie auch die Methoden zur Erfassung und die Ebenen (Einzelfall oder Aggregatenebene) der Zielerreichung bzw. der Wirkung liegen jedoch keineswegs und widerspruchsfrei vor, sondern müssen mit Verfahren der Qualitätsentwicklung und Evaluation erst bestimmt werden (vgl. Peters 2009, S. 168 ff.).

Pies und Schrapper (2001, S. 21) verweisen ferner darauf, dass es neben diesen politisch und ökonomisch motivierten Entwicklungen auch ein wachsendes Interesse der Profession an der empirischen Überprüfung und Bestätigung ihres Handelns bzw. der Wirkungen sozialpädagogischer Leistungen gibt, welches ebenfalls zur Erklärung der vermehrten Durchführung von entsprechenden Studien beiträgt. Hierzu zählen auch – zunehmend – Evaluationen. Ihre Bedeutung ergibt sich nahezu zwangsläufig aus dem Anspruch der Qualitätssicherung und -entwicklung. Anders als typische wissenschaftliche Studien, die eher an der Klärung allgemeiner Fragestellungen interessiert sind, bewerten Evaluationen – grob gesprochen – spezifische Maßnahmen hinsichtlich ihrer Zielsetzungen, Umsetzungen und Ergebnisse. Jugendämter, die im Kontext des NPM vermehrt als ‚Einkäufer‘ von Dienstleistungen auftreten, wollen wissen, was sie für welchen Preis einkaufen. Auch lässt sich ein vermehrtes Interesse der Träger sozialer Dienstleistungen an der Qualität der eigenen Leistungen ausmachen, um konkurrenzfähig zu bleiben. „Evaluation ist so untrennbar verbunden mit Qualitätssicherung und Entwicklung“ (Böttcher et al. 2006, S. 7). Genauer: „Der Zweck, für den die Akteure die Mühe der Evaluation auf sich nehmen, besteht primär darin, genauere Informationen zur Bewertung der Qualität der Arbeit und Hinweise für eine gezielte Verbesserung von Strukturen, Prozessen und Ergebnissen des Handelns zu erhalten“ (Merchel 2010, S. 18f.). Auch wenn in der Praxis Evaluationen nicht immer in ein Qualitätsmanagement eingebunden sind, ein Qualitätsmanagement kommt ohne Evaluation nicht aus.

Zum genaueren Verständnis des Gesagten müssen wir uns kurz mit dem Konzept Evaluation befassen.

Exkurs: Evaluation – was ist das?

Pragmatisch betrachtet lässt sich Evaluation wie folgt definieren: Evaluation ist ein wissenschaftsbasiertes Bewertungsverfahren. Der häufigste Gegenstand von Evaluationen sind Programme. Unter einem Programm versteht man im Kontext von Evaluation eine Maßnahme, in der auf Basis eines Konzepts planvoll Aktivitäten durchgeführt werden, die zu bestimmten Wirkungen für bestimmte Zielgruppen führen soll: z. B. Interventionen, Projekte, Modellversuche. Der Programm-Begriff steht gelegentlich auch für ein Bündel solcher Maßnahmen.

Evaluation ist so etwas wie ein Alltagsbegriff geworden, ein Containerbegriff zudem. Vieles wird Evaluation genannt, das mit Bewerten und Werten zu tun hat: Meinungsäußerungen, der schlichteste Feedbackbogen, Audits, Inspektionen, Tests, Leistungsmessungen oder Controllingverfahren. Bereits diese Skizze macht deutlich, dass Evaluationen sich von anderen Verfahren der systematischen Bewertung unterscheiden. „Feedbacks“ folgen zwar Regeln, sind aber eher subjektive Rückmeldungen mit lediglich hinweisendem Charakter, „Monitoring“ meint ein Beobachtungsverfahren, das Indikatoren sammelt, die Aussagen über bestimmte Erfolge eines Systems treffen können, „Audits“ überprüfen, ob bestimmte und wohldefinierte Verfahren eingehalten wurden, „Tests“ sind dazu angelegt, Kompetenzen von individuellen Akteuren zu überprüfen, Controlling hingegen zielt auf Planung, Steuerung und Kontrolle ab.

Zwei zentrale Merkmale also kennzeichnen Evaluation: Sie liefert – erstens – ein Werturteil. Eine Bewertung ist zwingend, soll es sich um Evaluation handeln. Um – zweitens – zu diesem Urteil zu kommen, folgt sie den Regeln wissenschaftlicher Forschung. Die Bewertung basiert also auf theoretischen Annahmen, einem systematischen Methodeneinsatz, der Ermittlung von Evidenzen (Belege, Indizien, robuste Hinweise, Beweismittel) und logischen Schlussfolgerungen. Eine Evaluation muss selbstverständlich den wissenschaftlichen Gütekriterien der Validität, der Reliabilität und der intersubjektiven Überprüfbarkeit genügen. Soll eine Bewertung erfolgen, benötigt man Kriterien, anhand de-

rer sie erfolgt. Dass diese in der Praxis der Forschung durchaus unterschiedlich sein können, zeigen die in dieser Einführung vorgestellten Studien eindrücklich.

Basisinfo: Wissenschaftliche Gütekriterien

Validität, Reliabilität und Intersubjektivität gelten als Gütekriterien wissenschaftlicher Forschung: Validität gibt an, inwieweit mit einer Messung (z. B. einem Test oder einer Befragung) das gemessen wird, was die Messung angibt zu messen. Reliabilität meint die Qualität eines Messinstrumentes, Ergebnisse zuverlässig zu reproduzieren. Intersubjektivität (oder Objektivität) meint die Unabhängigkeit der Messung von der die Messung durchführenden Person(en).

Diaz-Bone weist darauf hin, dass es problematisch wäre, die Gütekriterien nur auf den Messprozess zu reduzieren, da „die Qualität der Forschung sich als Resultat der Qualität des gesamten Forschungsprozesses ergibt“ (2015, S. 169).

Wichtig ist auch die Unterscheidung von interner und externer Validität. Erstere liegt vor, wenn Veränderungen der abhängigen Variablen auf den Einfluss der unabhängigen beruhen. Externe Validität meint: Man kann das Ergebnis einer Stichprobe über diese Stichprobe hinaus generalisieren.

Zu den grundlegenden Merkmalen von Evaluation zählt drittens ihr Dienstleistungscharakter. Eine Evaluation hat in aller Regel eine_n Auftraggeber_in, jemand, der ein bestimmtes Interesse an einer Evaluation hat. „Der zentrale Punkt dabei ist: Evaluation soll einen Nutzen haben. Sie erfüllt keinen Selbstzweck“ (Widmer/De Rocchi 2012, S. 11). Zu den Zwecken einer Evaluation können zwei wesentliche gezählt werden: Es geht um Legitimation, um eine mögliche Modifikation des Gegenstands, der evaluiert wird oder um eine Entscheidungsfindung: ja oder nein.

Wenn wir aus Evaluationssicht über HzE reden, dann haben wir es in aller Regel mit Maßnahmen bzw. Interventionen zu tun, also mit Aktivitäten, die ein Ziel oder mehrere aufeinander bezogene Ziele verfolgen – also es geht um „Programm-Evaluation“.

Die Ziele sozialer – und pädagogischer – Programme beziehen sich auf die Förderung von Erziehung, Lernen und Kompetenzentwicklung, sie sollen vulnerable Personen schützen, Beratung oder Hilfe verbessern oder die Bedingungen für diese Aktivitäten sichern. Ebenso können die Organisationen, in denen diese Programme eingebettet sind, Gegenstand sein. Organisationen sind soziale Systeme, die sich durch Mitgliedschaft, Dauerhaftigkeit, Strukturierung und Zielorientierung kennzeichnen lassen, in den HzE insbesondere die öffentlichen und freien Träger der Kinder- und Jugendhilfe.

Ein paar letzte Hinweise zur begrifflichen Orientierung: Mit Evaluation kann sowohl die Aktivität des Evaluierens als auch das Ergebnis dieser Aktivität gemeint sein – oder eben undifferenziert beides. In der Regel wird durch den jeweiligen Kontext der Verwendung klar, was gemeint ist. Manchmal wird auch von Evaluationsforschung als Synonym von Evaluation gesprochen. Womöglich soll hierdurch die wissenschaftliche Basis der Evaluation verdeutlicht werden. Andere verstehen unter Evaluationsforschung die Forschung über Evaluation. Wir schließen uns der letzteren Verwendung an, Verwechslungen können vermieden werden, wenn man auf den Begriff Evaluationsforschung ganz verzichtet.

Auch die Unterscheidung von Grundlagenforschung, angewandter Forschung und Evaluation ist nicht einfach zu treffen, denn diese drei Verfahren kommen in der Realität nicht in „Reinheit“ vor. Im Hinblick auf Methodologie und Methoden weisen sie keine Differenzen auf. Wissenschaftliche Forschung wie auch Evaluation nutzen die üblichen Verfahren empirischer Forschung. Es gibt quantitative wie qualitative Zugänge. Methodologische Überzeugungen, die Anwendung unterschiedlicher Instrumente, die Art und Weise, wie die mit diesen Instrumenten erhobenen Daten ausgewertet und interpretiert werden, all das lässt sich nicht dem einen oder anderen Ansatz zuordnen. Am deutlichsten werden die Unterschiede im Vergleich zwischen Grundlagenforschung und Evaluation. Im Kern geht es um die Fragen, wie spezifisch versus allgemein oder wie nutzenorientiert versus geleitet durch „reines“ Forschungsinteresse gearbeitet

wird. Evaluation bewertet also ein bestimmtes Programm (z. B. die Hilfe zur Erziehung, die der Träger XY in seinem neuen ambulanten Projekt anbietet, Forschung untersucht den Typus „ambulante HzE“). Evaluation will z. B. dem Träger vermitteln, wie er das Projekt verbessern kann, Forschung hingegen will zum wissenschaftlichen Fortschritt beitragen, indem z. B. Mechanismen der Wirkung aufgedeckt werden. Auch hinsichtlich des Ursprungs der beforschten Fragen kann der Unterscheid deutlich werden: In der Forschung ist es nicht unüblich, dass der/die Forscher_in selbst ein Frageinteresse hatte. Um finanziert zu werden, muss er gegebenenfalls Geldgeber überzeugen. Die typische Evaluation hingegen hat eine_n Auftraggeber_in, der/die an den Ergebnissen interessiert ist. Schließlich spielt das Bewerten die entscheidende Rolle: Während das Werten in der (Grundlagen)forschung (ideal gesprochen) „verboten“ ist, kann man nicht von Evaluation sprechen, wenn nicht bewertet wird. Welche Bedeutung dabei die entsprechenden Kriterien und damit implizit oder explizit die dem Phänomen (hier erzieherische Hilfen) zugewiesenen Funktionen oder Ziele haben, zeigen die unterschiedlichen Studien in Kapitel 2 eindrücklich.

Kurz muss noch erläutert werden, was es für die Evaluation bedeutet, wenn hier von Programmen gesprochen wird. Rossi/Lipsey/Freeman zeigen in ihrer klassischen Einführung in die Evaluation pädagogischer oder sozialer Programme fünf relevante Fragedimensionen auf (zum Folgenden vgl. Rossi/Lipsey/Freeman 2004, S. 53 ff.): das ‚Needs Assessment‘, das ‚Assessment of Program Theory‘, das ‚Assessment of Program Process‘, das ‚Impact Assessment‘ und schließlich das ‚Efficiency Assessment‘. Das soll knapp erläutert werden.

1. Die Bewertung des Bedarfs (Needs Assessment): Ein Programm zielt primär darauf ab, ein Problem lindern zu wollen. Um den genauen Bedarf zu ermitteln, müssen Art und Umfang des Problems erfasst, die Notwendigkeit für eine Intervention und die Gestaltung der Intervention bestimmt werden. Insbesondere muss hier entschieden werden, welche Interventionen erforderlich sind und ob es bereits verlässliche

Programme gibt. Im Prinzip geht es hier um Fragestellungen, die vor der Implementierung eines Programms gestellt werden müssen. Die radikalste: Benötigen wir das Programm überhaupt?

2. Die Bewertung der Programmtheorie (Assessment of Program Theory): Ein Problem und den Interventionsbedarf erkannt zu haben, bedeutet nicht, dass jede diesbezügliche Intervention funktioniert. Daher müssen die Konzeptionen und Designs von Programmen auf eine gültige Theoriebildung hin untersucht werden. Diese sollte schlüssige Annahmen darüber enthalten, wie und warum das Programm funktionieren sollte. Diese Annahmen müssen zu einer Programm-Theorie ausgearbeitet werden.
3. Die Bewertung der Programm-Prozesse (Assessment of Program Process): Auch wenn Programmtheorie und Design auf einen günstigen Programmverlauf hindeuten, so stellt sich jedoch häufig in der Praxis heraus, dass Programme nicht gemäß ihrer Vorgaben implementiert werden. Die angemessene Durchführung eines Programms wird häufig durch die Rahmenbedingungen, falsche Methodenwahl oder durch wenig kooperative Programmteilnehmer_innen oder -durchführende behindert. Die Zuverlässigkeit der Leistungserbringung (durch die Durchführenden), die Wahl der richtigen Adressat_innen, die effektvolle Organisation und Ressourcennutzung usw. stehen hier im Mittelpunkt. Prozessevaluation ist die häufigste Form einer Programmevaluation. Sie wird entweder als eigenständige Evaluation oder in Verbindung mit einer Outcomeevaluation als Teil einer umfassenden Evaluation genutzt.
4. Die Bewertung der Wirkungen (Impact Assessment): Die Bewertung des Outcomes gibt Antwort darauf, in welchem Maß ein Programm die intendierten Verbesserungen der Bedingungen und Ergebnisse erzielt und ob die Veränderungen auch nichtintendierte Nebenwirkungen generieren. Die Bewertung des Outcomes ist der stringenteste Weg, um zu beurteilen, ob ein Programm die angesteuerten Effekte bewirkt.

5. Bewertung der Programmeffizienz (Efficiency Assessment): Selbst wenn sich ein Programm als äußerst wirksam erweist, kann es auf Grund einer ineffizienten Kosten-Nutzen-Relation wenig sinnvoll sein, es erneut durchzuführen. Daher muss sich Evaluation auch mit der Frage befassen, inwiefern die Programmleistungen in einem angemessenen Verhältnis zu den Programmkosten stehen.

Dieser schnelle Blick auf Dimensionen von Evaluationen weist durchaus bestimmte Überschneidungen mit der oben zitierten Trias der Struktur-, Prozess- und Ergebnisdimensionen von Programmen auf. Spezifisch für Evaluation ist die Frage nach ihrem Verwendungsanspruch und den Verwertungsmöglichkeiten ihrer Ergebnisse. Grundsätzlich dienen Evaluationen dazu, Programme (aber eben auch Organisationen) datenbasiert zu beschreiben, zu bewerten und zur Verbesserung von Problemlösungen beizutragen (vgl. Heiner 2001, S. 483). Damit ergibt sich, dass Evaluation sowohl auf Ebene der Profession (also zu allererst für die sozialpädagogischen Akteure), als auch auf Ebene der politischen Steuerung relevant ist.

Von entscheidender Bedeutung sind bei jeder Evaluation die Adressat_innen und Nutzer_innen der Programme. Die Ergebnisse von Evaluationen sind relevant für die Adressat_innen der sozialpädagogischen Interventionen, wie auch für die Professionellen. Aber „hinter“ den pädagogischen Akteuren stehen in aller Regel auch deren Auftraggeber und Geldgeber – und mit Blick auf die Soziale Arbeit am Ende entsprechende gesellschaftliche Erwartungen bzw. Interessen. Auch die für politisch-strategisches Handeln zuständigen Personen oder Personengruppen können Adressat_innen der Befunde sein. Selbstverständlich ist die Evaluation auch ihren Auftraggeber_innen gegenüber verpflichtet. Auftraggeber kann im Prinzip jede der eben genannten Gruppen sein. Es kann sich aber auch um einen weiteren Akteur handeln. Es ist üblich geworden, alle, die „irgendwie“ an Evaluation beteiligt oder von ihr „betroffen“ sind, als „Stakeholder“ zu bezeichnen. Um diesen Begriff frei zu übersetzen: Diese Akteure haben „Aktien im Spiel“, sie haben eigene Interessen. Wessen Interessen