

KONZEPTE DER WIRKUNGS- MESSUNG UND -FORSCHUNG | Zwischen Goldstandard und vergoldeten Attrappen

Armin Schneider

Zusammenfassung | Wirkung ist ein Begriff, mit dem sich die Komplexität in Non-Profit-Organisationen nur bedingt erfassen oder messen lässt. Unterschiedliche Konzepte der Evidenzbasierung, des Wirkungscontrollings und der qualitativen Wirkungsforschung werden diskutiert. Neuere Konzepte wie das CAIMeR-Modell und der Einbezug von qualitativen Evidenzen können für die Soziale Arbeit dienlicher sein als rein quantitative Wirkungskonzepte, die entgegen ihrer Behauptung eher vergoldete Attrappen als ein Goldstandard sind.

Abstract | The term "effect" has limitations for the description or measurement of the complexity in non-profit-organisations. Different concepts of Evidence-based Practices, of Effect-Controlling and qualitative research of effects are discussed in this contribution. New concepts like the CAIMeR-Model and the inclusion of qualitative evidences are much more appropriate for Social Work as pure quantitative concepts, which are in fact against their own proposition rather gold-plated dummies than gold-standard.

Schlüsselwörter ▶ Wirksamkeitsforschung
▶ Qualität ▶ Evidenz ▶ Soziale Arbeit
▶ Non-Profit-Organisation

Einleitung | Wirkung ist im Grunde ein Konzept, das aus den Naturwissenschaften kommt und relativ linear von einer Ursache auf eine Wirkung schließen lässt. Dieses Wirkungskonzept hat in den letzten Jahren Einzug in die Arbeit von Non-Profit-Organisationen gehalten getreu dem Motto: Wo viel Geld reinfließt, müssen auch Wirkungen erzielt werden, die es zu messen gilt. Der Sozialen Arbeit als Profession wie auch anderen Non-Profit-Bereichen ist es nach ihrem eigenen Verständnis wichtig, für Individuen, Gruppen, Organisationen und Gesellschaften etwas zu bewirken, etwas besser zu machen, Partizipationschancen zu erhöhen, persönliches Wachstum zu fördern und die Handlungsoptionen zu erweitern. Zunehmend

stellt sich dabei die Frage, was davon im Sinne einer Ursache-Wirkung-Beziehung messbar ist, was sich einer direkten Messbarkeit entzieht oder gar nicht messbar ist.

Zunächst muss kurz untersucht werden, was den Gegenstand Sozialer Arbeit und dessen Besonderheiten ausmacht, bevor dann darauf eingegangen wird, mit welchen Konzepten eine Wirkungsforschung oder gar eine Wirkungsmessung überhaupt beziehungsweise mit welchen Chancen, Risiken und Grenzen sie eine Funktion für die Soziale Arbeit und damit für die ihr anvertrauten Personen erfüllen kann.

Zum Wirkungsbegriff | Das Wort Wirkung hat seinen Ursprung auch im Werken und kann als beabsichtigte Veränderung beschrieben und definiert werden. Gerade diese Veränderung ist im sozialwissenschaftlichen Bereich jedoch nicht einfach nachzuvollziehen, zu beobachten oder gar zu messen. Eine Wirkung in der Sozialen Arbeit wie auch in Non-Profit-Organisationen kann zum einen selten direkt beobachtet werden, zum anderen bedarf die Messung einer Wirkung mindestens zweier Zeitpunkte, schließlich lassen sich externe Effekte nur sehr schwer „herausrechnen“. Wenn ein Kind in einer Kindertagesstätte an einer Sprachförderung teilnimmt, so können zum Beispiel externe Effekte (neue Freunde, eine veränderte Umgebung, ein Entwicklungssprung, eine andere Bezugserzieherin etc.) stärker positiv oder negativ auf die Sprachkompetenz des Kindes wirken, so dass die Wirkung der intendierten Sprachförderung weder eindeutig zu beschreiben noch zu berechnen ist. Ob und wie und in welchem Zeitraum eine Organisation wie Greenpeace mit einer Kampagne die Umwelt nachhaltig schützt oder das Bewusstsein der Menschen beeinflusst, lässt sich ebenso wenig direkt und eindeutig messen.

Stockmann und Meyer nennen drei Wirkungsdimensionen, die bei einer Wirkung zu berücksichtigen sind: zum einen die Dimension von Struktur, Prozess und Verhalten Einzelner oder von Organisationen, dann die Dimension des „geplant“ oder „ungeplant“ und schließlich die Dimension positiv oder negativ (Stockmann; Meyer 2014, S. 78). Soll eine Wirkung gemessen werden, beispielsweise in Form einer Wirkungsevaluation, so lassen sich Bruttowirkungen (alle erkennbaren Effekte, das heißt inklusive externer Effekte) und Nettowirkungen (nur die auf die Inter-

vention zurückzuführenden Effekte) unterscheiden. Die Bruttowirkungen setzen sich entsprechend der Wirkungsformel (*ebd.*, S. 79) aus den Interventions-effekten (Nettowirkungen), externen anderen Faktoren (externe Faktoren) und Designeffekten zusammen. Letztere entstehen allein durch die Tatsache, dass eine Untersuchung stattfindet, sozusagen durch das Messen und die Beobachtung (vgl. auch Hawthorne-Effekt).

Wirkungsformel Bruttowirkung = Nettowirkung (Interventionseffekte) + externe Faktoren + Designeffekte (*Stockmann; Meyer 2014, S. 78*)

Beispiel Soziale Arbeit: Individuen und Konstellationen eher komplex als einfach |

Gegenstände Sozialer Arbeit sind Komplexitäten und kaum standardisierbare Fakten. Daher ist eine einfache Messbarkeit selten gegeben beziehungsweise stets die Gefahr vorhanden, dass unwesentliche Aspekte gemessen werden. *Graßhoff (2012)* kritisiert die rein auf die Outcomes von Sozialer Arbeit reduzierten Formen der Wirkungsforschung: „Da diese Arbeiten fast ausschließlich nach dem Ursache-Wirkungsprinzip funktionieren, irritiert diese Zielrichtung die (sozial-)pädagogische Forscherin bzw. den Forscher. In einer Hierarchie von absteigender Güte empirischer Untersuchungen bilden die quantitativ-klinischen Modelle die absolute Spitze und qualitativ-rekonstruktive nur eine vorwissenschaftliche und maximal explorative Vorstufe. Diese Logik der Fetischisierung von positivistischer und deduktiv-nomologischer Forschung ist nicht nur aus normativen Gründen, sondern auch methodologisch und gegenstandsbezogen problematisch“ (*ebd.*, S. 73 f.).

Wenn Komplexitäten von Situationen von Individuen, Gruppen und Gesellschaft auf rein messbare Größen und nur auf erwiesene und nachweisbare Wirkungen reduziert werden, hat dies auch „Wirkungen“ auf die Form und Organisation Sozialer Arbeit: Nicht von der Hand weisen lässt sich der von *Ziegler* ins Feld geführte Hintergrund der Wirkungsforschung: „Ohne Zweifel folgt die derzeitige Wirkungsforschung im Wesentlichen utilitaristischen Begründungen und dient nicht selten als ein bloßes Informationsbeschaffungswerkzeug für managerielle Steuerungsfantasien“ (*Ziegler 2012, S. 94*) oder mit *Rock* ausgedrückt: „Je nachdem, ob und inwieweit man sich auf Wirkungsorientierung einlässt, verändert das auch den Charak-

Gegenläufig

Im April war es wieder soweit: Mit der Veröffentlichung des Deutschen Freiwilligensurveys 2014 wird unserem Land der engagementpolitische Spiegel vorgehalten. Seit 1999 gibt das Bundesfamilienministerium diese Studie alle fünf Jahre heraus, nun erstmals unter der wissenschaftlichen Leitung des DZA – Deutsches Zentrum für Altersfragen, Berlin. Mit seinen 644 Seiten ist der Survey ein Füllhorn wissenschaftlich fundierter Fakten und Trends.

Die wichtigste Nachricht: 43,6 Prozent der Bevölkerung ab 14 Jahren waren im Jahr 2014 freiwillig engagiert, knapp zehn Prozentpunkte mehr als bei der ersten Erhebung im Jahr 1999 (34 Prozent). Die Spenderquote ist im selben Zeitraum aber nahezu stetig gesunken: von 63,4 Prozent (1999) auf nur noch 54,4 Prozent im Jahr 2014. Wie lässt sich diese gegenläufige Entwicklung von Zeitspende und Geldspende erklären? Der Survey nennt gleich mehrere mögliche Gründe: Entsolidarisierung (auch als Folge abnehmender religiöser Orientierung), Vertrauensverlust in Spendenorganisationen, demgegenüber die unmittelbare Glaubwürdigkeit und Transparenz eigenen ehrenamtlichen Engagements auf lokaler Ebene, und schließlich ein steigender Anteil von Personen, die sich aufgrund ihrer knappen finanziellen Ressourcen nicht mehr in der Lage sehen, Geld zu spenden.

Für gemeinnützige Organisationen bestätigt der neue Freiwilligensurvey Herausforderungen, denen sich die meisten von ihnen schon seit Längerem bewusst sind. Es wird schwieriger, die Bevölkerung zum Spenden zu motivieren. Die Schaffung von Transparenz und Vertrauen erfordert außerdem einen immer größeren Ressourceneinsatz; das vorliegende Schwerpunktthema zur Wirkungsmessung zeigt dies auf bezeichnende Weise. Dass sich diese Anstrengungen aber durchaus lohnen, ist zum Beispiel an der positiven Entwicklung der 30 größten Organisationen mit DZI Spenden-Siegel ablesbar: Ihr Geldspendenvolumen stieg von 840 Mio. Euro (2000) auf 900 Mio. Euro (2009) und 1,1 Mrd. Euro 2014.

Burkhard Wilke
wilke@dzi.de

ter der Sozialen Arbeit. Die Wirkungsfrage wird so zu einem Instrument der fiskalischen und bürokratischen Selbstkontrolle“ (Rock 2015, S. 173).

Bleck und Liebig (2015, S. 167 f.) sehen die Begriffe Qualität, Wirkung und Nutzen in einem Zusammenhang: Während Qualität im Sinne einer Messung und eines Managements eher die institutionelle Sicht mit quantitativen Daten darstellt und der Nutzen eher im Sinne einer Nutzer- und Nutzungsforschung die Sicht der Nutzerinnen und Nutzer mit qualitativen Daten, steht die Wirkung als Evaluations- und Wirkungsforschung in der Mitte zwischen den beiden Begriffen der Qualität und des Nutzens. Für die Praxis bedeutet dies: „Angeraten wird also ein Diskussionsprozess, in dem geklärt wird, zu welchen Phasen der konkreten Erbringungskontexte überhaupt Wissen zu der Ergebnisqualität, den Wirkungen oder dem Nutzen der Dienstleistungsprozesse für die Organisationen und für deren Fachlichkeit, deren Wettbewerbsfähigkeit, deren strategische Ausrichtung etc. hilfreich sein kann. In idealer Weise finden in dieser Planungsphase alle hier gegeneinander gestellten Konzepte Berücksichtigung – auch wenn in der Umsetzungsphase sicherlich nicht alle in gleichem Maße Anwendung finden“ (ebd., S. 168).

Eindeutig muss von der Disziplin Sozialer Arbeit die Frage beantwortet werden, ob und wie Wirkungsforschung und -konzepte ihrem Gegenstand und Selbstverständnis und ihrer Wirkung gerecht werden. Wie sich später herausstellen wird, gibt es beispielsweise mit dem CAImeR-Modell bereits eigene Anstrengungen aus der Sozialen Arbeit, welche die Komplexität besser abbilden als reine Ursache-Wirkungs-Modelle. Diese Konzepte können für die Wirkungsdiskussion auch in anderen Non-Profit-Bereichen wichtige Impulse geben, denn auch dort kommen einfache quantitative Beschreibungen an ihre Grenzen.

Was bedeutet messen? | Selbst in der quantitativen Forschung wird deutlich, dass Merkmale oder Merkmalsausprägungen zwar beobachtet, aber nicht direkt gemessen werden können; die Übersetzung in messbare Größen bietet die Variable. Das Merkmal schlecht zum Beispiel ist beobachtbar, aber als solches nicht messbar. Erst durch den Kunstgriff, aus dem Merkmal die Variable zu machen, wird zum Beispiel das Geschlecht eine Variable mit verschiedenen Ausprägungen, denen Zahlen zugeordnet werden

können. Dabei ist es für das Messen unerheblich, ob der Ausprägung „weiblich“ die Zahl 1 oder 796 zugeordnet wird. Die Schwierigkeit der Messung besteht darin, dass die Zuordnung von Zahlen auch aussagekräftig sein muss (Döring; Bortz 2016, S. 235). Demnach kann jede Messung an dem Punkt kritisiert werden, an dem Messfehler entstehen können, oder einfacher ausgedrückt: Die Messung stellt immer nur eine verkürzte Abbildung der Wirklichkeit dar, wie auch eine Landkarte nur einige Dimensionen einer tatsächlichen Landschaft darstellt und andere bewusst auslässt oder gar nicht darstellen kann.

Bei der Messung von Wirkungen muss diese Grundtatsache der Messbarkeit, die zum einen nur eine Abbildung ist und zum anderen immer mehr oder weniger große Messfehler beinhaltet, berücksichtigt werden. In der Messtheorie werden vor allem drei Probleme dargestellt, darunter das Repräsentationsproblem, die Frage also danach, ob ein Merkmal überhaupt messbar ist. Weitere Probleme sind das Eindeutigkeitsproblem und das Bedeutsamkeitsproblem. Mathematisch bedeutet Ersteres, dass Messwerte nur insoweit verändert werden dürfen, dass die Aussagen dabei nicht verändert werden. Das heißt beispielsweise, dass Aussagen einer nominalen Skala nicht zu einer Rangordnung führen dürfen: „Im Zuge der statistischen Auswertung dürfen nur die jeweils zulässigen Transformationen durchgeführt werden, damit die inhaltliche Aussage der Messwerte erhalten bleibt“ (Döring; Bortz 2016, S. 236). Das Bedeutsamkeitsproblem wirft die Frage nach der Sinnhaftigkeit der Messungen und deren Ergebnisse für den Untersuchungsgegenstand auf.

Auch wenn die genannten Probleme Berücksichtigung finden, kann jede Messung nur eine Annäherung an die Wirklichkeit sein. Zu unterscheiden sind bei den Konzepten der Wirkungsmessung und -forschung vor allem drei unterschiedliche Vorgehensweisen, zwischen denen naturgemäß viele Formen liegen:

- ▲ Messungen von ausschließlich messbaren Dimensionen (rein quantitative Vorgehensweisen);
- ▲ Ablehnung von Messungen, stattdessen Beschreibungen (rein qualitative Vorgehensweisen);
- ▲ sich gegenseitig ergänzende Konzepte von Messung und Beschreibung unter Nutzung verschiedener Verfahren (einander ergänzende Methoden der Wirkungsmessung).

„Was tun? Anstatt sich auf einen Rechenwettbewerb einzulassen, gilt es, die bearbeiteten sozialen Probleme, die spezifische Methode der Organisation, die handelnden Personen und die Erfolge gemeinnütziger Organisationen transparent darzustellen. Diese vier Merkmale sind das „magische Viereck“ für die Darstellung der Leistung. Wer sich daran orientiert, der generiert mehr Vertrauen als es eine SROI-Kalkulation vermag. Mit Recht heißt es: Nicht alles, was gezählt werden kann, zählt. Und nicht alles, was zählt, kann gezählt werden“ (Rock 2015, S. 175).

Konzepte der Messbarkeit | Der Social Return on Investment (SROI) versucht ebenso wie der Social Reporting Standard soziale Aus- und Nebenwirkungen auf die Gesellschaft messbar zu machen. Beim Social Reporting Standard werden Struktur- und Prozessqualität sowie intendierte soziale Wirkungen beschrieben (Schellberg 2015, S. 116). Beim SROI in seiner ursprünglichen Form wurden die Wirkungen vor allem in Geldströmen gemessen (ebd., S. 118). Der auf den europäischen Kontext angewendete SROI der xit GmbH besteht aus sechs verschiedenen Perspektiven der Wirkung (ebd., S. 120):

- ▲ institutionelle Transferanalyse: Geldfluss zwischen der Institution und der öffentlichen Hand;
- ▲ individuelle Transferanalyse: Geldfluss zwischen Leistungsempfängern und der öffentlichen Hand;
- ▲ Alternativbetrachtung/Opportunitätsenerträge: Welche Kosten und Erträge hätte die öffentliche Hand, wenn es die Sozialleistung nicht gäbe;
- ▲ regionalökonomische Wirkung: Nutzen der regionalen Wirtschaft durch die Institution;
- ▲ Wirkungen auf die Lebensqualität der Leistungsempfänger;
- ▲ Wirkungen auf die gesellschaftliche Wohlfahrt/das Sozialklima.

Insbesondere die letzten beiden Bereiche zeigen, dass nicht alle Felder mit Geldströmen messbar sind. Für diese Bereiche werden folgende Indikatoren einer Selbst- und Fremdeinschätzung unterzogen (Schellberg 2015, S. 135):

- ▲ emotionales Wohlbefinden;
- ▲ soziale Beziehungen;
- ▲ materielles Wohlbefinden;
- ▲ persönliche Entwicklung;
- ▲ physisches Wohlbefinden;
- ▲ Selbstwirksamkeit;
- ▲ gesellschaftliche Teilhabe;
- ▲ Rechte.

Die Autoren des SROI warnen vor einem naturalistischen Fehlschluss und Entscheidungsautomatismen: „Menschen in den verschiedenen Kontexten müssen am Ende ethisch wertende Entscheidungen treffen – Sozialarbeiter in konkreten Einzelfällen, Geschäftsführer und Einrichtungsleiter für ihre Organisationen, Politiker in politischen Gremien. Doch die Entscheidungsrationalität kann sich mit dem Social Return on Investment deutlich verbessern“ (ebd., S. 137).

In ähnlicher Weise beschreibt die New Economics Foundation die unterschiedlichen gesellschaftlichen Beiträge verschiedener Professionen für die Gesellschaft. Hier wurden sechs Professionen untersucht und folgende hervorsteckende Ergebnisse erzielt: „While collecting salaries of between £500000 and £10 million, leading City bankers to destroy £7 of social value for every pound in value they generate [...] For every £1 they are paid, childcare workers generate between £7 and £9.50 worth of benefits to society“ (New Economics Foundation 2009, S. 3). Bei dieser Untersuchung wurde ebenfalls ein SROI-Konzept zugrunde gelegt. Hierbei standen Auswirkungen des Handelns der einzelnen Professionen insbesondere auf einzelne Stakeholder und ausschließlich materielle Werte im Vordergrund (ebd., S. 29).

Evidence-based Practice | Evidenzbasierte Ansätze, ob in der Sozialen Arbeit, in der Medizin oder im Management, gehen von einer Messbarkeit aus und beschreiben Ansätze, die davon geleitet sind, dass es eindeutige und richtige Methoden für eine Behandlung oder Beratung gibt. Dabei stützen sie sich auf eine Reihe von empirischen Daten. Unter dem Motto und der pragmatischen Zielsetzung „What works“ stehen Wirkungen im Vordergrund bis hin zu der Maxime, dass nur eine Praxis, deren Evidenz eindeutig nachgewiesen ist, auch Berücksichtigung finden kann. „This modern movement towards evidence-based practice is founded on the presupposition that the best way to improve outcomes in a range of applied settings is to identify empirical evidence that illustrates which practices are most effective [...] Evidence based practice has become dominant across many fields, and because qualitative research has tended to be relegated to the lower end of the evidence hierarchy pyramid [...] there is the risk of less qualitative evidence being produced by the qualitative community“ (O'Reilly; Kiyimba 2015, S. 167).

Für die Evidence-based Practice (EBP), die seit 1997 in den USA große Bedeutung hat, steht die grundsätzliche Frage im Vordergrund, worauf sich Entscheidungen in der Sozialen Arbeit stützen, was hier Wirksamkeit bedeutet und ob und ein positiver Effekt auf die Adressatinnen und Adressaten zu beobachten ist (James 2015). Wurde in den Anfängen und gerade von europäischer und deutscher Seite zu Recht die Konzentration auf den sogenannten Goldstandard der RCTs (Randomized Controlled Trials) kritisiert (Otto u.a. 2009, Sommerfeld; Hüttemann 2007), so werden inzwischen stärker Konzepte der Triangulation und der Mixed Methods eingesetzt, um Evidenz nicht ausschließlich auf quantitativ erhobene Fakten zu stützen. Deutlich wird, dass gerade in der Sozialen Arbeit die Arbeit nach dem Prinzip von Versuch und Irrtum nicht hinreichend ist. Aus Sicht von Befürwortern hat Evidence-based Practice es erst ermöglicht, Wissensbestände in der Sozialen Arbeit umzusetzen (James 2015). Als fallbezogener Prozess stehen vor allem drei Pole, aus der Evidence-based Medicine kommend, für die Evidence-based Practice: Auf der einen Seite die beste Evidenz aus der Forschung, die individuelle klinische Expertise und die Präferenzen und Werte des Patienten. Damit ist EBP mehr als nur die Anwendung und Umsetzung von durch Forschung generiertem Wissen. Für den Prozess bedeutet dies (ebd.):

- ▲ Problemformulierung;
- ▲ systematische Suche;
- ▲ kritische Bewertung;
- ▲ Einschätzung mit Adressatenpartizipation sowie
- ▲ Evaluation und Feedback.

In den Sozialwissenschaften, insbesondere in der Sozialen Arbeit, sprechen eine Reihe von Gründen für eine EBP, die versucht, Forschungsergebnisse wohlfundiert und allgemeingültig zu nutzen. Gefragt werden muss, ob erstens eine Recherche in jedem Einzelfall aufwendig ist und, zweitens, nur eine schwach entwickelte Forschungsbasis in der Sozialen Arbeit vorhanden ist (wenig Wissen über effektive Praktiken, die Mehrheit von evaluativen Studien hat nur schwach entwickelte Forschungsdesigns; James 2015). Erst das Interesse der Nutzerinnen und Nutzer im angelsächsischen Raum machte auf eine Reihe von Praktiken aufmerksam (Liste der Prävention, Triple P, KEEP etc.). In diesem Kontext sind auch mehr oder weniger abgesicherte Verfahren entstanden, wie zum Beispiel die Kompetenzorientierung als Methodik in der Kinder- und Jugendhilfe (Cassée 2010).

Als Kernelemente der EBP nennt James (2015) am Beispiel jugendlicher Delinquenz:

- ▲ reflektiertes, systematisches Denken über Risiko- und Schutzfaktoren der Delinquenz;
- ▲ Integration von diversen forschungsbasierten Handlungsmethoden auf allen Ebenen des Umfelds;
- ▲ unterstützende Trainings- und Supervisionsstrukturen;
- ▲ ausgeprägte Elternarbeit;
- ▲ Stärkenorientierung;
- ▲ Kooperation und Interdisziplinarität;
- ▲ Manualisierung: Arbeiten nach Handlungsschemata, die mehr sind als Standardisierung, theoretische und methodische Aufschlüsselung sowie Validität.

Aber: Es gibt nicht nur Dienstleistungen, die das „Gütesiegel“ einer EBP haben!

Bei der Implementation ist die Identifikation von Praktiken nur der erste Schritt. Zu beachten ist eine translationale Agenda (knowledge transfer, knowledge utilization, dissemination, implementation, „from bench to trench“). Schließlich sind Kontexte der zeitlichen Gegebenheiten, der finanziellen Ressourcen, der Kapazitäten, der Einstellung des Personals, des politischen Willens, professioneller Werte, des Leadership und der Unterstützung durch das Kernpersonal bedeutsam. Mehr und mehr wird auch die EBP als iterativer Prozess verstanden, der mit einer Fallanalyse beginnt (James 2015). Durch EBP kann die Wissensbasis Sozialer Arbeit gestärkt werden, es werden konzeptionelle, empirisch verfechtbare Handlungsmethoden mit verbesserten Resultaten etabliert, es entstehen Praxis-Forschungs-Allianzen und eine erhöhte Professionalisierung (ebd.).

Nutley (2011) sieht verschiedene Möglichkeiten der Einbindung von Forschungsergebnissen in (politische) Entscheidungsprozesse und stellt die EBP damit in einem weiteren Kontext:

- ▲ Research-based Practitioner: Forschung im Dialog mit Praktikerinnen und Praktikern unter deren Verantwortung;
- ▲ Embedded-research Modell: Forschung ist eingebunden in Prozesse, Politik, Werkzeuge; Forschung als Instrument, Verantwortung liegt beim Management;
- ▲ Organisational-excellence Modell: lokale Anpassung, Organisation als Schlüssel, lokale, forschungsoffene Kultur, angemessene Dienstleistungskulturen, -strukturen und -prozesse.

Mullen (2014) sieht Grenzen in der rationalistischen Tradition der Evidence-based Practice und fügt den üblichen fünf Schritten der Evidence-based Practice (Ask – Acquire – Appraise – Apply – Analyse und Adjust) noch zwei weitere Schritte als Schritte 4 und 7 hinzu: *Relevance*, *Credibility* und *Strength* sowie *Reformulate evidential argument* und *Assess client system ongoing*. Als Quellen der Erkenntnis fügt er neben der Rezeption von Studien auch die Erkenntnisse der ermöglichenden und unterstützenden Variablen im „dort“ und „hier“ (horizontale Erforschung) und die Erkenntnisse aus der vertikalen Erforschung über die kausalen Prinzipien „hier“ und „dort“ hinzu. Ihm kommt es entscheidend auf die jeweiligen Bedingungen sowie die Passung mit dem konkreten System vor Ort an. Als erweiterte Methode stellt Mullen das CER-System vor (Comparative Effectiveness Research). Hierbei spielen neben den quantitativen experimentellen Studien auch qualitative Erkenntnisse und unterschiedliche Methoden eine Rolle. In ähnlicher Weise haben sich auch Otto, Polutta und Ziegler (2009, S. 248) für eine „zweite Generation“ der Evidence-based Practice im Sinne eines Evidence-based Reflexive Professionalism ausgesprochen, die sich aus unterschiedlichen Wissensbeständen speist.

Zusammengefasst kann daher von einer Entwicklung des Evidence-based-Practice-Modells hin zu einem umfassenderen Verständnis einer Evidence gesprochen werden, insbesondere durch eine deutlichere Berücksichtigung auch von Erkenntnissen, die nicht im Sinne eines quantitativ-standardisierten Forschungsparadigmas gewonnen werden.

Wirkungsorientiertes Controlling in Non-Profit-Organisationen | Auch im Controlling als betriebswirtschaftlicher Funktion der Information und Unterstützung der Leitung und Steuerung einer Organisation und eines Unternehmens durch Kennzahlen hat sich inzwischen die Erkenntnis durchgesetzt, dass der Erfolg insbesondere von Non-Profit-Organisationen nicht allein durch das Verhältnis von Aufwand und Ertrag (Output beziehungsweise Effizienz) beschrieben oder beurteilt werden kann. Im Modell der *International Group of Controlling* (2010, S. 47) werden als Effektivitätskriterien Effects (objektiv messbare Wirkung), Impacts (subjektive wahrgenommene Wirkung, beispielsweise bei Klientinnen und Klienten) und Outcomes (gesellschaftliche Wirkungen) beschrieben. Dies jedoch nicht in einem

abstrakten, kontextlosen Raum, sondern individuell bezogen auf die einzelne Organisation, die einer Analyse von relevanten Umweltfaktoren und Spezifitäten unterzogen wird (zum Beispiel in den unterschiedlichen Ausprägungsgraden zwischen Haupt- und Ehrenamtlichkeit, der Dienstleistungs- und der Voicefunktion, der Autonomie beziehungsweise Fremdregulierung; *ebd.*, S. 33). Auch wird erkennbar, dass Komplexität umfassender abgebildet werden kann.

CAIMeR | Das CAIMeR-Modell versucht generell, verschiedene Wirkfaktoren Sozialer Arbeit ergänzend zu den reinen Ergebnissen aufzuzeigen. Den Autoren ging es bei der Entwicklung des Modells darum, die „Black Box“ der Wirkungen zu beschreiben (Blom; Morén 2010). CAIMeR steht als Akronym für Kontexte (Contexts), Akteurinnen und Akteure (Actors), Interventionen (Interventions), Mechanismen (Mechanisms) und Resultate (Results). Dabei wird eine Breite von Faktoren einbezogen, einige Bestandteile der einzelnen Bereiche sind (*ebd.*, S. 107-111):

▲ Kontexte: der Intervention, der Lebenswelten und -wirklichkeiten der Adressatinnen und Adressaten, kulturelle und soziale Kontexte, fördernde und hindernde Faktoren und Beziehungen auf verschiedenen Ebenen.

▲ Akteurinnen und Akteure: alle am Interaktionsprozess beteiligten Akteurinnen und Akteure mit ihren Ansichten und Perspektiven. Diese Gruppe umfasst nicht nur Einzelpersonen, sondern auch Gruppen und Gesellschaften.

▲ Interventionen: Hierbei wird der Blick nicht nur auf die Sozialarbeiterinnen und Sozialarbeiter, sondern auch auf die Adressatinnen und Adressaten sowie alle sonst involvierten Personen und Organisationen sowohl in Bezug auf die formalen als auch auf die informellen Interventionen gerichtet.

▲ Mechanismen: Unter Mechanismen werden alle sozialen, psychosozialen und psychologischen Wirkungen, Rückwirkungen und Nebenwirkungen verstanden.

▲ Resultate: Ergebnisse in unterschiedlichen Zeithorizonten, auch in Bezug auf die Umgebung, auf Prozesse und weitergehende Aspekte.

Qualitativ-rekonstruktive Wirkungsfor-
schtung | Schmitt (2011, S. 188) geht davon aus, dass in herkömmlichen quantitativen Forschungen schon zu Beginn der Forschung ein Wirkungsmodell vorhanden ist, „um davon Hypothesen abzuleiten, die empi-

risch falsifiziert werden können und damit das Modell bestätigen oder widerlegen“ (ebd.). Dem gegenüber steht die Rekonstruktion im Vordergrund einer qualitativen Wirkungsforschung: „In der Tradition qualitativer Forschung sind solche Wirkungsmodelle das Ergebnis einer entdeckenden Forschung, denn sie kann dazu beitragen, mögliche kausale Pfade zu erwünschten und unerwünschten Effekten zu rekonstruieren [...]“ (ebd.).

Wesentliche Möglichkeiten einer qualitativen (Wirkungs-)Forschung sind demnach:

- ▲ „die Entdeckung von Wirkungen;
- ▲ die Entdeckung von nicht intendierten Nebenwirkungen;
- ▲ die Entdeckung von relevanten moderierenden Kontextbedingungen;
- ▲ die Aggregation von Wirkungen, deren Mechanismen und Kontextbedingungen zu einem Modell, das weiteres Forschen anregt (quantitativ und qualitativ);
- ▲ die Rekonstruktion von Wertungen im Feld, welche die Kommunikation der Evaluationsergebnisse beeinflussen“ (Schmitt 2011, S. 189).

Eine Möglichkeit qualitativer Wirkungsforschung stellt die Metaphernanalyse dar, die sich metaphorischen Redeweisen widmet und die darin enthaltenen Deutungsmuster regelgeleitet rekonstruiert (Schmitt 2011, S. 191). Mit dieser Methode können „wirkungsbehindernde Kontexte, nicht-intendierte Nebenfolgen wie Ansätze für wirkende Kommunikation“ (ebd., S. 199) identifiziert werden. „Die Metaphernanalyse dient also der Rekonstruktion individueller und kultureller Deutungsmuster. Sie bietet Regeln für die Durchführung der Identifikation von Metaphern, aber auch für die Gewinnung von Interpretationen und Regeln für Grenzen der Verallgemeinerung und des Anspruchs“ (ebd., S. 194). Auch in den Ansätzen einer qualitativen Wirkungsforschung zeigt sich der Hinweis auf eine Einbeziehung und Nutzung auch quantitativer Methoden.

Zusammenfassung und Ausblick | Die beste Wirkung zu untersuchen und dann nur noch anzuwenden, scheint das Heilsversprechen der Evidence-based Practice zu sein, dabei werden aber andere Konzepte und Ideen der Wirkungsforschung und -messung und die Tatsache übersehen, dass sich nicht alles messen lässt. Aber auch das Nichtmessbare wirkt! So kann zusammenfassend herausgestellt werden, dass bei den

hier vorgestellten Ansätzen der Wirkungsforschung mehrere Aspekte einbezogen werden (müssen), um Wirkung tatsächlich erforschen zu können. Zum einen sind dies mehrere Perspektiven (zum Beispiel der Sozialarbeiterinnen und -arbeiter, der Adressatinnen und Adressaten sowie anderer Akteurinnen und Akteure), verschiedene Forschungsrichtungen (zum Beispiel quantitativ-standardisiert und qualitativ-rekonstruktiv), unterschiedliche Zeitpunkte (zum Beispiel vor einer Intervention, nach einer Intervention, in mehreren Zeitabständen) wie auch die dahinter liegenden Interessen, Absichten und Werthaltungen (zum Beispiel Kürzungsabsichten, bestmögliche Behandlung und Beratung, Vereinfachung komplexer Zusammenhänge).

Der kausale Wirkungsbegriff steht in einer anderen Tradition als der Begriff der Professionalität und das scheint gerade für Non-Profit-Bereiche und deren Entwicklung und Ausprägung von besonderer Relevanz zu sein: „Dem Wirkungskonzept liegt eine Kausalität zugrunde, dem Professionalitätskonzept eine wissenschaftsfundierte Beziehung zwischen Theorie und Praxis, die sich unter anderem an einem Habitus, einer ethischen Orientierung und einer solchen am Gemeinwohl festmacht“ (Schneider 2011, S. 28). Insbesondere aus der Profession heraus bedarf es eines eigenen Verständnisses von Wirkung: „Der Maßstab für eine Wirkung darf nicht dem Markt und sekundären Interessen überlassen werden, sondern muss sich an professionellen Werten ausrichten [...]“ (ebd.). Gerade professionelles Handeln wirft den Blick auf das Besondere und das Allgemeine: „Professionen bilden naturgemäß Routinen aus, die sich aber in ihrer Wirksamkeit hinterfragen lassen müssen. Professionen müssen immer offen sein für Fälle, die der ‚Lehrmeinung‘ entgegenstehen und ihr jeweiliges nichtstandardisiertes Arbeitsbündnis mit jedem besonderen Fall besonders gestalten“ (ebd.).

Wirkungsforschung hält Perspektiven für die Soziale Arbeit ebenso wie für andere Non-Profit-Bereiche bereit und kann auch im Sinne der ethischen Werte Sozialer Arbeit (und anderer gesellschaftlicher Bereiche und Funktionen) auf der Basis von Menschenwürde und sozialer Gerechtigkeit erfolgreich sein, wenn sie sich als Annäherung an die Wirklichkeit versteht und mehrere Perspektiven, unter anderem und vor allem auch die der Betroffenen zu mehreren Zeitpunkten einbezieht. Wirkungsforschung

kann nur dann wirken, wenn sie verlässliche und valide Ergebnisse eines von der Sozialen Arbeit heraus definierten komplexen Wirkungsbegriffes aufzeigt. Allerdings muss darauf hingewiesen werden, dass derzeit weder aus der Sozialen Arbeit heraus noch von deren Geldgebern ernst zu nehmend Anstrengungen unternommen werden, eine verlässliche und solide Wirkungsforschung zu betreiben. Nur diese könnte für die Soziale Arbeit und ihre Adressatinnen und Adressaten Türen öffnen und auf der Basis von wissenschaftlichen Erkenntnissen Interventionen nach hinreichenden Anamnesen und diagnostischen Prozessen so befördern, dass sie nach dem jeweiligen „State of the Art“ erfolgen und dennoch den einzelnen Problemlagen entsprechen. Schließlich wäre zu wünschen, dass eine solche Entwicklung auch durch systematische Evaluationen begleitet und erweitert würde, damit eine kontinuierliche Verbesserung erfolgen kann.

Gerade bei der Evidence-based Practice zeigt sich, dass nicht alles Gold ist, was glänzt, und der vermeintliche Goldstandard von RCTs sich auch als vergoldete Attrappe erweisen kann. Jedes Wirkungsforschungskonzept muss dahingehend geprüft werden, ob es hält, was es verspricht. Dies muss in erster Linie eine wertgebundene Prüfung sein, die sich an dem Menschenbild und den Werten der Sozialen Arbeit orientiert.

Professor Dr. Armin Schneider ist Sozialarbeiter, Pädagoge und Theologe. Er lehrt und forscht in den Bereichen empirische Sozialforschung und Sozialmanagement an der Hochschule Koblenz. Armin Schneider war von 2010 bis 2016 Sprecher der Sektion Forschung der Deutschen Gesellschaft für Soziale Arbeit und ist Direktor des Institutes für Bildung, Erziehung und Betreuung in der Kindheit, Rheinland-Pfalz. E-Mail: schneider@hs-koblenz.de

Literatur

- Bleck**, Christian; Liebig, Reinhard: Qualität, Wirkung, Nutzen. Diskussionszusammenhänge und Zugänge zu Resultaten Sozialer Arbeit. In: Blätter der Wohlfahrtspflege 5/2015, S. 163-169
- Blom**, Björn; Morén, Stefan: Explaining Social Work Practice – The CAImeR Theory. In: Journal of Social Work 1/2010, pp. 98-119
- Cassée**, Kitty: Kompetenzorientierung. Eine Methodik für die Kinder- und Jugendhilfe. Bern 2010

Döring, Nicola; Bortz, Jürgen: Forschungsmethoden und Evaluation. Berlin 2016

Graßhoff, Gunther: NutzerInnen und AdressatInnen als produktive Kraft des Sozialen. In: Standpunkt Sozial 1-2/2012, S. 69-75

International Group of Controlling: Wirkungsorientiertes NPO-Controlling. Leitlinien zur Zielfindung, Planung und Steuerung in gemeinnützigen Organisationen. Freiburg im Breisgau 2010

James, Sigrid: Inside the belly of the beast – Möglichkeiten und Grenzen der evidenzbasierten Praxis. Vortrag auf der DGSA-Jahrestagung in Würzburg am 25.4.2015, eigene Notizen des Autors

Mullen, Edward: Rethinking the Idea of Evidence in Evidence-based Policy & Practice. Vortrag am 16.4.2014 in Bozen, eigene Notizen des Autors

New Economics Foundation: A Bit Rich: Calculating the real value to society of different professions. London 2009

Nutley, Sandra: Evidence informed practice. Vortrag in Utrecht 2011

O'Reilly, Michelle; Kiyimba, Nikki: Advanced Qualitative Research. Los Angeles 2015

Otto, Hans-Uwe; Polutta, Andreas; Ziegler, Holger (eds.): Evidence based Practice – Modernising the Knowledge Base of Social Work? Opladen 2009

Rock, Joachim: Wirkung und Werte: zwei Seiten einer Medaille. In: Blätter der Wohlfahrtspflege 5/2015, S. 173-175

Schellberg, Klaus: Der Social Return on Investment: Strategische Möglichkeiten für den Sozialbereich? In: Sprinkart, Karl Peter (Hrsg.): Nachhaltigkeit messbar machen. Regensburg 2015, S. 113-137

Schmitt, Rudolf: (Nicht-)Wirkungen erkunden: Möglichkeiten und Grenzen der systematischen Metaphernanalyse in der sozialwissenschaftlichen Wirkungsforschung. In: Eppler, Natalie; Mieth, Ingrid; Schneider, Armin (Hrsg.): Qualitative und quantitative Wirkungsforschung. Opladen 2011, S. 185-202

Schneider, Armin: Professionelle Wirkung zwischen Standardisierung und Fallverstehen: Zum Stand der Wirkungsforschung. In: Eppler, Natalie; Mieth, Ingrid; Schneider, Armin (Hrsg.): Qualitative und quantitative Wirkungsforschung. Opladen 2011, S. 13-32

Sommerfeld, Peter; Hüttemann, Matthias: Evidenzbasierte Soziale Arbeit. Nutzung von Forschung in der Praxis. Baltmannsweiler 2007

Stockmann, Reinhard; Meyer, Wolfgang: Evaluation. Opladen 2014

Ziegler, Holger: Wirkungsforschung – über Allianzen on Evaluation und Managerialismus und die Möglichkeit erklärender Kritik. In: Schimpf, Elke; Stehr, Johannes (Hrsg.): Kritisches Forschen in der Sozialen Arbeit. Wiesbaden 2012, S. 93-105