

# WIRKUNGSANALYSE IN DER SOZIALEN ARBEIT | Differenzierung ist nötig

Sebastian Ottmann; Joachim König

**Zusammenfassung** | Die Debatte über Wirkungen Sozialer Arbeit gewinnt erneut an Dynamik. Im Beitrag wird vor allem diskutiert, wie Wirkungen analysiert und gemessen werden können. Eine differenzierte Auseinandersetzung zeigt hier, dass begrifflich und methodisch oft unscharf agiert wird und so auch die Interpretation der Befunde nicht selten problematisch erscheint. Deshalb ist ein differenzierter und methodisch seriöser Umgang mit dem Begriff Wirkung wichtig. Der Artikel präsentiert eine empirische Toolbox, die ein stufenweises und praxistaugliches Verfahren beschreibt und so eine differenzierte und empirisch solide Auseinandersetzung mit Wirkungen in der Praxis ermöglicht.

**Abstract** | Recently, the discourse on effects and impacts of social work gains speed. Main topic is how to analyze and measure effects. A critical approach proves that fuzziness often prevails, conceptually as well as methodically. Thus, the interpretation of relevant findings is often problematic. The concept of effect requires a differential and methodically reliable procedure. This article presents an empiric tool box for a graduated procedure which allows a sound analysis with effects in practice.

**Schlüsselwörter** ► Soziale Arbeit ► Wirkung  
► Messung ► Evaluation ► Methode

**Einleitung** | Nicht zuletzt durch das Bundesfreiheitsgesetz (BTHG) hat die Debatte über die Wirkungen der Sozialen Arbeit wieder an Dynamik gewonnen. So zeigt sich, dass in den letzten Jahren wieder vermehrt zum Thema publiziert wurde (Borrmann; Thiessen 2016 sowie die Ausgabe 6-7/2016 dieser Zeitschrift). Aber auch Entwicklungen wie der Social Return on Investment (Schober; Then 2015) oder die Entwicklung des Social Entrepreneurship geben dem Thema Wirkungsorientierung und Wirkungsanalyse in der Sozialen Arbeit neuen Aufwind. Allerdings zeigt sich bei der Diskussion auch, dass nicht immer methodisch seriös mit dem Begriff der Wirkung umgegan-

gen wird. Zentral bei diesem Begriff ist ein zugrunde liegendes kausales Verständnis. Dies zeigt sich auch bei der Definition von Wirkung im Duden, die diese als eine „durch eine verursachende Kraft bewirkte Veränderung“ (*Bibliographisches Institut o.J.*) definiert. Die verursachende Kraft ist in der Sozialen Arbeit die untersuchte Maßnahme beziehungsweise das untersuchte Angebot. Der kausale Mechanismus ist damit zentral und es darf nur von einer gefundenen Wirkung gesprochen werden, wenn diese mit sehr hoher Wahrscheinlichkeit auch auf die untersuchte soziale Intervention rückführbar ist. Dieser Anspruch wird nicht immer verfolgt, dennoch wird von einer gefundenen Wirkung gesprochen.

Neben dieser notwendigen und klaren Definition gibt es bei der Begriffsbestimmung von Wirkung weitere Punkte zu beachten:

- ▲ Die Wirkung (Outcome und Impact) muss ganz klar vom Output<sup>1</sup> abgegrenzt werden. Outcome nimmt die Wirkungen auf der Ebene der Klientinnen und Klienten in den Blick, während Impact die gesellschaftliche Wirkung beschreibt.
- ▲ Wirkung ist nicht Ergebnisqualität!<sup>2</sup> Vielfach wird dies in der Debatte vermischt, allerdings können nicht beide Begriffe synonym verwendet werden, auch wenn einzelne Indikatoren in beiden Bereichen zum Einsatz kommen können. Zentral bei der Analyse und Erfassung der Wirkung ist aber der oben beschriebene kausale Mechanismus, der bei der Messung von Ergebnisqualität keine Berücksichtigung findet.
- ▲ Eine Unterscheidung, die auch mit dem Kausalitätsverständnis zusammenhängt, ist die zwischen der Brutto- und Nettowirkung (Caspari 2009). Ziel von Wirkungsanalysen sollte immer die Erfassung der Nettowirkung sein, also der Wirkung, die auf die soziale Maßnahme zurückgeführt werden kann. Oft wird aber in empirischen Untersuchen die Bruttowirkung in den Blick genommen und es kann nicht ausgeschlossen werden, dass die erfasste Wirkung nicht durch andere Einflussfaktoren bedingt wurde. Unter der Bruttowirkung wird die Veränderung zwischen zwei Zeitpunkten verstanden, unabhängig davon, ob diese durch die untersuchte Maßnahme zustande gekommen ist oder nicht.

<sup>1</sup> Outputs werden „dingliche Resultate eines Programms, welche als Leistungen oder Produkte durch (oder in Verbindung mit den) Programm-Interventionen hervorgebracht werden“ (EvalWiki 2012), definiert.

<sup>2</sup> Im Rahmen des Qualitätsmanagements werden mit der Ergebnisqualität die Resultate eines Prozesses näher analysiert (Hensen 2019, S. 33).

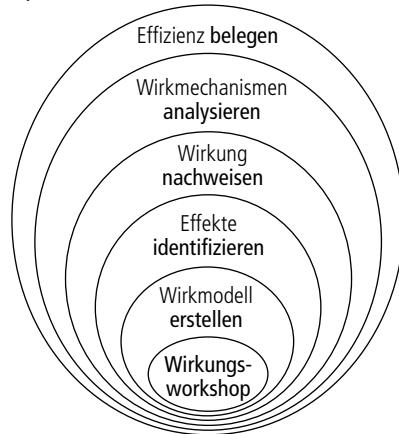
Um einen differenzierteren Umgang mit dem Begriff Wirkung leisten zu können und deutlich zu machen, dass eine Wirkung eben häufig nicht ohne Einschränkung kausal auf eine Maßnahme zurückgeführt werden kann, sollte mit dem Begriff der Wirkungsplausibilisierung gearbeitet werden, den Balzer (2012) einföhrte. Balzer versteht hierunter alle empirischen Designs, die keinen empirischen Wirkungsnachweis erbringen können, beispielsweise Veränderungsmessungen oder theoriebasierte Evaluationsansätze. Im Rahmen einer Wirkungsplausibilisierung sollte aber dann auch bei diesen Designs eine Einschätzung abgegeben werden, ob die gefundenen Veränderungen auf die Maßnahme zurückzuführen sind beziehungsweise welche weiteren Einflussfaktoren es für diese gibt oder geben könnte.

Die Definition des Wirkungsbegriffs hat auch Auswirkungen auf die Wahl des empirischen Designs bei Wirkungsanalysen und vor allem bei der Berichterstattung der Ergebnisse. Hierbei sollte nicht immer vorschnell von Wirkungen gesprochen werden, vielmehr muss auch eine Beurteilung über die Aussagekraft der durchgeföhrten Untersuchung und deren Einschränkungen erfolgen. Dies ist vor allem auch wichtig, um die Akzeptanz und Glaubwürdigkeit von Wirkungsanalysen im Feld der Sozialen Arbeit zu erhöhen, zumindest aber, um sie nicht zu gefährden.

### **Der Wirkungsradar als empirische Toolbox |**

Der Wirkungsradar (siehe Abbildung 1), der am Institut für Praxisforschung und Evaluation der Evangelischen Hochschule Nürnberg entwickelt wurde, stellt eine empirische Toolbox dar, mit der ein differenzierter Umgang mit Wirkungsanalysen in der Sozialen Arbeit möglich ist. Vor allem klar abgegrenzte, präventive und intervenierende Maßnahmen und Programme können damit untersucht werden. Bei stationären Angeboten oder Gesamtplanverfahren wird der Forschungsgegenstand jedoch in der Regel komplexer und es muss gut überlegt werden, welches Vorgehen hier gewählt wird. Beim Einsatz des Wirkungsradars in einer solchen Untersuchung müssen in einem ersten Schritt die zentralen Angebote des Settings identifiziert werden. Danach kann der Wirkungsradar auf diese einzelnen Angebote angewendet werden, um mögliche Effekte und Wirkungen zu identifizieren. Nach diesen Analysen wird abschließend eine übergreifende Untersuchung durchgeföhr, in der die Koordination und die Stimmigkeit der verschiedenen

**Abbildung 1: Der Wirkungsradar als empirische Toolbox für Wirkungsanalysen (Ottmann; König 2018)**

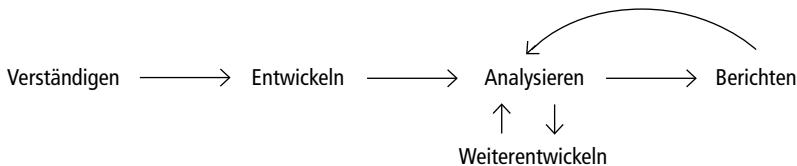


Angebote in den Blick genommen werden. Durch dieses Vorgehen können dann auch valide Aussagen über die Effekte und die Wirksamkeit von stationären Einrichtungen beziehungsweise Gesamtplanverfahren erzielt werden.

Bevor die Toolbox näher beschrieben wird, muss aber noch kurz darauf eingegangen werden, dass die eigentliche Wirkungsanalyse nur einen Schritt im Rahmen der Wirkungsorientierung darstellt (Abbildung 2). Erst wenn innerhalb einer Einrichtung oder eines Trägers ein Verständnis von Wirkung und Wirkungsorientierung erarbeitet wurde und Angebote sowie Maßnahmen wirkungsorientiert entwickelt wurden, ist eine Wirkungsanalyse sinnvoll. Anschließend müssen die gefundenen Ergebnisse in der Einrichtung mit möglichst vielen Mitarbeitenden interpretiert, diskutiert und hinsichtlich der fachlichen Weiterentwicklung daraufhin überprüft werden, ob an bestehenden Konzeptionen Überarbeitungen beziehungsweise Ergänzungen nötig sind. Hier wird in den nächsten Jahren die Herausforderung innerhalb der Sozialen Arbeit liegen, Kompetenzen in Einrichtungen, bei Trägern und Wohlfahrtsverbänden aufzubauen, um einen solchen Prozess zwischen Entwicklung, Analyse, Weiterentwicklung und Berichterstattung regelmäßig zu durchlaufen.

Der Wirkungsradar greift daher auch die zwei Schritte im Implementierungskreislauf vor der eigentlichen Wirkungsanalyse auf und beginnt bei der Erarbeitung eines allgemeinen Verständnisses zum Thema.

**Abbildung 2: Implementierungskreislauf von Wirkungsorientierung (Ottmann; König 2018)**



Bevor die eigentlichen Schritte näher dargestellt werden, soll aber noch kurz auf die Grundannahmen des Instrumentariums eingegangen werden (Ottmann; König 2018, S. 14 ff.):

- ▲ ein stufenweises Vorgehen und die Grundlage einer fachlichen Perspektive im Rahmen der Wirkungsanalyse;
- ▲ ein partizipativer Ansatz, bei dem die Fach- und Führungskräfte aus der Praxis mit der Wissenschaft zusammenarbeiten. Dadurch kann sichergestellt werden, dass die Erfahrungen und das Wissen der Praxis in den Prozess und die Analysen einfließen können. Die beteiligten Wissenschaftlerinnen und Wissenschaftler stellen die Einhaltung wissenschaftlicher Standards und den Einsatz adäquater Methoden im Prozess sicher;
- ▲ die Verwendung von sogenannten Black-Box- (Frage, ob ein Angebot wirkt) und White-Box-Verfahren (Frage, wie die Wirkung zustande kommt);
- ▲ das Konzept der Realistic Evaluation als forschungsmethodisches Grundverständnis mit dem Ziel der Entwicklung einer Theorie mittlerer Reichweite.

Die Anwendung des Konzeptes der Realistic Evaluation von Pawson und Tilley (1997) macht vor allem in den Arbeitsfeldern der Sozialen Arbeit Sinn, da diese von einem linearen Kausalitätsverständnis abweicht. Nach den Autoren entsteht eine Wirkung (Outcome) immer aus dem Zusammenspiel von Kontext und Mechanismus, die in sogenannten KMO-Konfigurationen dargestellt werden. Es wird daher folgende Logik einer realistischen Erklärung zugrunde gelegt: „Ein Programm beziehungsweise eine Intervention erzeugt nur dann Wirkungen, wenn diese in einem konkreten Kontext Wirkungszusammenhänge aktivieren, die sich beobachten lassen und Veränderungen erzeugen“ (Haunberger; Baumgartner 2017, S. 126). Dieses Verständnis scheint gerade in den Arbeitsfeldern der Sozialen Arbeit sinnvoll, da bei Maßnahmen und Angeboten in diesem Bereich eine Vielzahl von Außeneinflüssen und Kontextfaktoren die Wirkung bei Klientinnen und Klienten mit verur-

sachthaben könnte. Auch erscheint das Ziel der Entwicklung von Theorien mittlerer Reichweite (Merton 1968, Pawson 2010), die „keinen Anspruch auf Generalisierung [haben], [...] aber über Aussagen zu einem Einzelfall hinaus[gehen]“ (Haunberger; Baumgartner 2017, S. 125), als sinnvolle Zielvorgabe für eine Wirkungsanalyse. Dies bedeutet aber auch, dass Erkenntnisse nicht automatisch verallgemeinert werden können und beispielsweise gefundene Wirkungen durch eine Replikation der Maßnahme an einem anderen Standort nicht automatisch eintreten müssen.

Mit dem Wirkungsradar werden in erster Linie die Wirkungen auf der Ebene der Klientinnen und Klienten einer sozialen Dienstleistung in den Blick genommen. Natürlich können die dargestellten Methoden und Schritte auch bei der Analyse von gesellschaftlichen Wirkungen (Impact) Anwendung finden, allerdings stellt sich hier die Frage, ob solche durch eine einzelne Maßnahme valide nachgewiesen werden können.

In einem ersten Schritt wird ein *Wirkungsworkshop* durchgeführt, um ein einheitliches Verständnis von Wirkung in der Einrichtung beziehungsweise beim Träger zu erarbeiten. Im Rahmen des Workshops sollten neben Führungskräften auch Mitarbeiterinnen und Mitarbeiter, die täglich im Kontakt mit Klientinnen und Klienten stehen, teilnehmen. Das erarbeitete Verständnis dient zum einen als Grundlage für die weiterführenden Analysen, kann aber auch in die aktuelle Diskussion mit Kostenträgern oder der Politik eingebracht werden.

**Entwicklung von Wirkmodellen** | Eine weitere wichtige Grundlage für die darauf aufbauenden empirischen Analysen ist die *Entwicklung von Wirkmodellen*. Ein Wirkmodell kann man als „eine logische, grafisch aufbereitete Darstellung dessen, wie ein Programm [oder eine Maßnahme] unter gewissen Rahmenbedingungen theoretisch und/oder empirisch funktioniert“ (Rauscher u.a. 2015, S. 43), verstehen. Ein solches Wirkmodell sollte im Rahmen der Wirkungsori-

tierung für jedes neue Angebot entwickelt werden, aber auch für bestehende Angebote, für die eine Wirkungsanalyse durchgeführt wird. Dieses kann auch für sich genommen gegenüber dem Kostenträger ein wichtiges strategisches Instrument sein, um die eigene Arbeit und die zu erzielenden Wirkungen transparent darzustellen (Ottmann; König 2019). Die Erstellung eines Wirkmodell sollte in Workshops erfolgen, an denen acht bis zehn Personen teilnehmen. Neben Führungskräften und Mitarbeitenden der Einrichtung können dies auch Klienten und Klientinnen vertretende Akteure sein, welche die Sicht der Klientinnen und Klienten einbringen können. Als Vorarbeiten empfiehlt es sich, bestehende Konzepte zu sichten, aber auch im Rahmen einer Literaturrecherche zu überprüfen, ob es schon Veröffentlichungen über mögliche Wirkungen und Wirkmechanismen gibt. Auch können qualitative Interviews mit Experten und Expertinnen sowie Klientinnen und Klienten durchgeführt werden, wenn mögliche Mechanismen noch erarbeitet werden müssen.

Als Methode zur Erstellung eines Wirkmodells hat sich die „Theory of Change“ als sehr praktikabel erwiesen (The Center for Theory of Change o.J.). Bei dieser wird zuerst die Hauptwirkung einer Maßnahme definiert, um danach in einer Rückwärtsinduktion die Teilwirkungen festzulegen, die erzielt werden müssen, damit die Hauptwirkung eintreten kann. Hierbei empfiehlt sich eine Orientierung an der Resultatetreppen (Beywl; Niestroj 2009), die zwischen Outcomes im Sinne einer Aneignung von Wissen und Fähigkeiten einerseits und der Veränderung von Handlungen und Verhalten im nächsten Schritt unterscheidet. In einem dritten Schritt im Rahmen der Theory of Change werden die Annahmen über den Kontext beschrieben, um danach festzulegen, welche Aktivitäten für die Erzielung der Hauptwirkung benötigt werden. In einem letzten Schritt werden bereits bei der Erstellung des Wirkmodells Indikatoren für eine spätere Überprüfung festgelegt. Am Ende des Prozesses wird das erstellte Wirkmodell visualisiert und verschriftlicht.

Wichtig bei der Erstellung ist zu überprüfen, welche Zielgruppen durch das Angebot erreicht werden und ob für diese unterschiedliche Wirkmodelle erarbeitet werden müssen (Schmidt 2016, S. 171 ff.). Auch sollte ein verstärktes Augenmerk auf Kontextfaktoren und mögliche Außeneinflüsse gerichtet werden. Hierbei sollte zwischen einrichtungsspezifischen Struktu-

ren, Inputs, die ein Träger in die Maßnahme einbringt, Ressourcen, die Klientinnen und Klienten einbringen (Incomes), sowie dem Kontext, der sich aufgrund gesellschaftlicher, politischer und rechtlicher Rahmenbedingungen ergibt, unterschieden werden (Bartsch u.a. 2016). Kontextfaktoren, die auch einen Einfluss auf die zu erzielende Wirkung auf der Ebene der Klientinnen und Klienten haben, sollten als Außeninflüsse gekennzeichnet werden. Auch muss festgehalten werden, ob mit einem positiven oder negativen Einfluss gerechnet wird und wie stark der Einfluss auf die Klientinnen und Klienten sowie die Maßnahme ist (Ottmann; König 2019).

Im Rahmen des Wirkungsradars werden Kontextfaktoren und Außeneinflüsse im Wirkmodell dargestellt und fließen dann – sofern sie erfasst werden können – als Störvariablen beim kausalen empirischen Nachweis der Wirkung ein. Im Rahmen des Schrittes „Wirkmechanismen analysieren“ kann dann überprüft werden, welchen Einfluss die Kontextfaktoren bei der Erzielung der Wirkung haben. Hierbei kann eine Verbindung zur Implementationsforschung hergestellt werden, die unter anderem die Implementationsqualität von wirksamen Programmen untersucht (Unz; Taube 2016). Ein Wirkmodell sollte im Rahmen eines Workshops immer konsensual beschlossen und mögliche alternative Erklärungen ausgeschlossen werden. Gerade letzteres – also das ständige Mitdenken von Alternativen – ist besonders wichtig, um einen sogenannten „confirmation bias“ (Vaessen 2016) zu verhindern. Unter diesem Begriff wird verstanden, dass die Ergebnisse einer Wirkungsanalyse ein Wirkmodell bestätigen, weil der Blick nur auf dieses eine Modell gerichtet wird.

**Identifizierung von Effekten** | Der dritte Schritt im Wirkungsradar, die *Identifizierung von Effekten*, ist ein zentrales Element für jede soziale Einrichtung, die eine wirkungsorientierte Arbeitsweise verfolgt, und sollte daher nachhaltig in der Einrichtung implementiert werden. Um Effekte, also Veränderungen bei Klientinnen und Klienten in Bezug auf die durchgeführte Maßnahme im Sinne der zuvor definierten (Outcome-)Ziele identifizieren zu können, muss in der Einrichtung ein wirkungsorientiertes Monitoring eingerichtet werden. Unter einem Monitoring versteht man die „Erhebung von Daten zu aufeinander folgenden Zeitpunkten zu gleichen Merkmalen eines Programms“ (EvalWiki 2015). Es wird also ein Längs-

schnittdesign implementiert, bei dem es mindestens einen Messzeitpunkt am Anfang und einen Messzeitpunkt am Ende der Maßnahme geben muss, um die Veränderungen erfassen zu können. Die Implementierung eines solchen wirkungsorientierten Monitorings ist auch deswegen unabdingbar, weil damit frühzeitig mögliche negative und unerwünschte Wirkungen der Maßnahme ersichtlich werden und der Träger dann entsprechend konzeptionell reagieren kann.

Das Erhebungsinstrument sollte die zentralen Output- und Outcome-Indikatoren in den Blick nehmen. Weiterhin können im Rahmen des Monitorings auch die Aktivitäten, die Inputs sowie soziodemografische Merkmale der Klientinnen und Klienten miterfasst werden. Gerade Letztere werden bei der Analyse von Wirkmechanismen benötigt. Wird ein solches wirkungsorientiertes Monitoring implementiert, sollte überprüft werden, inwieweit die Erhebung in bestehende Dokumentationsprozesse integriert werden kann, um den Mehraufwand der Mitarbeiterinnen und Mitarbeiter so gering wie möglich zu halten. Auch muss eine Entscheidung getroffen werden, ob die Erhebung im Rahmen einer Selbsteinschätzung (die Klientinnen und Klienten füllen die Erhebungsinstrumente selbst aus) oder im Rahmen einer Fremdeinschätzung beispielsweise durch die Fachkräfte erfolgt. Bei beiden Verfahren sind Verzerrungseffekte nicht ausgeschlossen. Nach allen Erfahrungen in der Praxis hat sich gezeigt, dass in den Angeboten der Sozialen Arbeit eine Fremdeinschätzung durch die Fachkraft oft zu validieren Ergebnissen kommt. Sinnvoll ist hierbei, die Bewertung mit den Klientinnen und Klienten zu besprechen, damit eine konsensuale Bewertung erfolgt.

Im Rahmen der Datenauswertung findet die Darstellung von Veränderungen bei den Klientinnen und Klienten statt. Neben der Auswertung über alle Teilnehmenden an einer Maßnahme können auch Extremgruppen (zum Beispiel die Personen mit den besten und den schlechtesten Veränderungen) oder Gruppen nach soziodemografischen Merkmalen (zum Beispiel nach Geschlecht, Alter oder Zugangsweg) in der Analyse näher betrachtet werden. In diesem Schritt wird auch die oben genannte Wirkungsplausibilisierung durchgeführt. In einem Interpretationsworkshop wird diskutiert, inwieweit die gefundenen Effekte auf die Maßnahme zurückzuführen sind beziehungsweise welche weiteren Einflüsse diese bedingen.

**Kausaler Wirknachweis |** Im nächsten Schritt findet dann der eigentliche *kausale Wirknachweis* statt. Um diesen realisieren zu können, ist neben den Personen, die an der Maßnahme teilgenommen haben, eine Vergleichsgruppe unabdingbar. In einer Vergleichsgruppe werden Personen befragt, die den Personen in der Maßnahmengruppe ähnlich sind, an der eigentlichen Maßnahme aber nicht teilgenommen haben. In der Praxis hat sich die Arbeit mit bestehenden Gruppen – im Rahmen eines quasi-experimentellen Designs – am handhabbarsten erwiesen. Dass ein solches Design zum kausalen Nachweis von Wirkungen nötig ist, zeigt auch die Evidenzhierarchie, die unter anderem von Veerman und van Yperen (2007) vorgestellt wurde. In dieser Hierarchie wird das Experiment als beste Methode zum Nachweis herausgestellt, danach folgen die quasi-experimentellen Designs.

Um dies umsetzen zu können, eignet sich vor allem das Verfahren des Propensity Score Matching (Rosenbaum; Rubin 1983), mit dem die Zusammenstellung der Vergleichsgruppe erfolgen kann. Im Rahmen dieses Verfahrens werden zentrale soziodemografische Merkmale und Störvariablen berücksichtigt und mit diesen wird der sogenannte Propensity Score berechnet. Im Rahmen des Matchings erfolgt dann der Abgleich zwischen der Maßnahmengruppe und der Vergleichsgruppe. Jeder Person in der Maßnahmengruppe wird ein Teilnehmer, eine Teilnehmerin aus der Vergleichsgruppe zugeordnet, und zwar die Person, die einen möglichst ähnlichen Propensity-Score-Wert aufweist. Wurde der Analyseschritt durchgeführt, kann danach betrachtet werden, wie die Veränderung in den zentralen Outcome-Indikatoren ausfällt. Subtrahiert man in einem ersten Schritt zuerst die Mittelwerte beider Gruppen am Ende der Maßnahme und danach die Mittelwerte am Anfang der Maßnahme, kann man von diesen beiden Werten nochmals die Differenz bilden und erhält somit den kausalen Effekt der Maßnahme, den sogenannten Differenz-in-Differenz-Ansatz (Angrist; Pischke 2009, S. 227 ff.). Nachdem zuvor das Matching durchgeführt wurde und sich die Personen in beiden Gruppen in den ausgewählten Variablen sehr ähnlich sind, können diese Variablen als Störfaktoren ausgeschlossen werden und mit sehr hoher Wahrscheinlichkeit kann eine Rückführung der gefundenen Effekte auf die untersuchte Maßnahme erfolgen. In der Praxis zeigt sich allerdings, dass in vielen Arbeitsfeldern der Sozialen Arbeit die Bildung einer Vergleichsgruppe nicht im-

mer möglich ist. Sollte dies der Fall sein, kann geprüft werden, ob andere Verfahren wie zum Beispiel das Wartekontrollgruppen-Design (Kolip; Greif 2016) angewandt werden könnte. Sollte dies auch nicht der Fall sein, muss bei der Darstellung der Ergebnisse klar kommuniziert werden, dass keine kausale Rückführung der gefundenen Effekte auf die Maßnahme erfolgen kann.

Bei der Diskussion von Vergleichsgruppendesigns innerhalb der Sozialen Arbeit wird als Gegenargument oft angeführt, dass die soziale Wirklichkeit so vielfältig sei, dass nie alle Störvariablen berücksichtigt werden können. Dies ist sicherlich der Fall, trotzdem sollte beim Versuch, einen Wirkungsnachweis zu erbringen, der Anspruch vorhanden sein, alle denkbaren Störvariablen abzubilden und – sofern möglich – ein quasi-experimentelles Design zu realisieren.

**Wirkmechanismen analysieren** | Neben der Frage, ob eine Maßnahme wirkt, ist auch die Frage, wie eine Maßnahme wirkt, von zentraler Bedeutung. Diese Frage wird im fünften Schritt des Wirkungsradars in den Blick genommen, wenn *Wirkmechanismen analysiert* werden. In diesem Schritt werden die zuvor bei der Erstellung definierten Mechanismen mittels des sogenannten Pattern Matching empirisch überprüft (Giel 2013, S. 227 ff.). Hierbei kann man zwischen sozialen Mechanismen und Prozessmechanismen unterscheiden. Bei sozialen Mechanismen werden Veränderungen auf der Ebene der Klientinnen und Klienten in den Blick genommen, also ein kleiner Teil des Wirkmodells betrachtet. Demgegenüber werden bei der Analyse von Prozessmechanismen die definierten Ursachen-Wirkungsbeziehungen, also ein größerer Teil des Wirkmodells, untersucht (Schmitt

2018, S. 2). Um diese Analysen durchführen zu können, stehen vielfältige empirische Methoden zur Verfügung. Neben quantitativen Verfahren wie Mediationsanalysen und Strukturgleichungsmodelle sind dies vor allem auch qualitative Methoden, mit denen bestimmte Aspekte nochmals tiefergehend betrachtet werden können. Hierfür können beispielsweise Experteninterviews oder Fokusgruppen mit Mitarbeitenden und Klientinnen und Klienten durchgeführt werden, aber auch neuere Methoden, wie das Process Tracing (Beach; Pedersen 2013) zum Einsatz kommen, eine Methode, die die Identifizierung von kausalen Mechanismen zum Ziel hat.

Um diesen Schritt durchführen zu können und Mechanismen genauer zu analysieren, muss nicht unbedingt ein kausaler Wirkungsnachweis erfolgen. Vielmehr kann man diesen Schritt auch schon durchführen, wenn man Effekte identifiziert hat, und hier überprüfen, wie diese Zustände kommen und ob beispielsweise auch alle Personen in der Maßnahmenzielgruppe gleichermaßen von diesen profitieren. Auch ohne den kausalen Wirknachweis können die Ergebnisse einer solchen Analyse wichtige Anhaltspunkte für die fachlich-konzeptionelle Weiterentwicklung liefern.

In einem letzten Schritt im Wirkungsradar kann die *Effizienz einer Maßnahme* betrachtet werden. Diese Betrachtung wurde bewusst als letzter Analyseschritt aufgenommen, da hier nun die Perspektive gewechselt wird. Es steht nicht mehr unbedingt die fachliche Perspektive im Mittelpunkt, sondern die monetäre Bewertung eines Angebotes. Um die Effizienz zu belegen, wird versucht, die erfassten Wirkungen zu monetarisieren. Danach kann dieser Wert in das

## 40 Jahre dokumentierte Fachdiskussion 40 Jahre Sozialwissenschaftliche Literaturdokumentation

- Onlinezugang in über 200 Hoch- und Fachhochschulbibliotheken
- Individuelle Beratung und Recherche mit Dokumentenlieferung

Deutsches Zentralinstitut für soziale Fragen/DZI

[www.dzi.de](http://www.dzi.de)

DZI SoLit

Verhältnis zu den eingesetzten finanziellen Mitteln gesetzt werden. Durch diese Berechnung wird der sogenannte Social Return on Investment (Schöber; Then 2015) ermittelt, der in diesem Schritt des Wirkungsradars die Standardmethode ist. Um eine belastbare und valide Bewertung der Effizienz durchführen zu können, sollte im Rahmen der Berechnung nur eine nachvollziehbare Monetarisierung betrachtet werden, das heißt diese sollte naheliegend und zum Zeitpunkt der Analyse belegbar sein. Auch sollten finanzielle Effekte, die in einer fernen Zukunft, etwa in drei bis fünf Jahren liegen, nicht berücksichtigt werden. Die valide Bewertung der Effizienz kann dann im Rahmen dieses Schrittes auch durch die interdisziplinäre Zusammenarbeit mit Wirtschaftswissenschaftlerinnen und -wissenschaftlern und Gesundheitsökonomien und -ökonominnen sichergestellt werden.

In diesem Vorgehen kommen verschiedene Stufen des sozialwissenschaftlichen Forschungsprozesses zur Geltung. So werden mit der Entwicklung des Wirkmodells Hypothesen generiert, beispielsweise über mögliche Wirkmechanismen, die dann in den nachfolgenden Schritten überprüft werden können. Dazu werden die einzelnen Schritte des Forschungsprozesses durchlaufen, die in der jeweiligen Stufe immer wieder mit dem entsprechenden Forschungsziel umgesetzt werden. Bei den Schritten „Effekte identifizieren“ und „Wirkung nachweisen“ werden quantitative Methoden eingesetzt, während bei den Schritten „Wirkmodelle entwickeln“ und „Wirkmechanismen analysieren“ auch qualitative Methoden zum Einsatz kommen.

Wurde eine Wirkungsanalyse mit dem Wirkungsradar durchgeführt, kann zur Darstellung der Ergebnisse der Social Reporting Standard (*Social Reporting Initiative e.V.* 2014) angewendet werden. Der Social Reporting Standard ist ein standardisiertes Berichtsformat für Programme und Maßnahmen im Non-Profit-Bereich, bei dem die Darstellung der erzielten Wirkungen eine wichtige Rolle spielt. Dieser Standard ist daher als Reporting-Instrument für Analysen mit dem Wirkungsradar gut eignet. Neben diesem Format, das sich zur Berichterstattung vor allem gegenüber Stakeholdern gut eignet, sollten die Ergebnisse aber auch für die weitere Nutzung im Rahmen der konzeptionellen Arbeit in der Einrichtung beziehungsweise innerhalb des Trägers aufbereitet werden. Dies kann beispielsweise in Form von Dashboards geschehen,

mit denen die Mitarbeitenden selbst Analysen abrufen können beziehungsweise die Ergebnisse interaktiv analysieren können (zum Beispiel im Rahmen von Gruppenauswertungen). Ein solches Dashboard empfiehlt sich vor allem für das im dritten Schritt implementierte Monitoring, da dann immer eine Aktualisierung mit den neu erhobenen Daten erfolgen kann. Im Hinblick auf die fachliche Arbeit empfiehlt es sich, die gefundenen Ergebnisse und zentrale Wirkmechanismen als Infografik oder Infoflyer aufzubereiten, um eine verständliche und einfache Beschreibung dieser oft komplexen Zusammenhänge für die Praxis zu ermöglichen (Macsenare 2018, S. 31).

Im Rahmen des Beitrags wurden methodische Mindeststandards beschrieben, die nötig sind, um die Wirkung einer sozialen Maßnahme kausal nachzuweisen. Zu beachten ist allerdings, dass auch experimentelle und quasi-experimentelle Methoden nur eingeschränkt aussagekräftig sein können. Diese Einschränkungen stehen sehr häufig im Zusammenhang mit der internen und externen Validität der Untersuchung.<sup>3</sup> Die externe Validität beschreibt, inwiefern es möglich ist, „die Ergebnisse der Evaluationsuntersuchung auf andere Kontexte zu generalisieren“ (Gollwitzer; Jäger 2014, S. 42). Die externe Validität spielt eine wichtige Rolle, da hierdurch nochmals der Kontext einer Maßnahme betont wird. Ein Wirkungsnachweis bedeutet damit nicht, dass eine Maßnahme ohne Weiteres an einem anderen Standort übernommen werden kann und dann die gleichen Wirkungen erzielt werden. Vielmehr empfiehlt es sich, eine Maßnahme unter verschiedenen Kontextbedingungen im Rahmen einer Wirkungsanalyse zu überprüfen, um eine Generalisierbarkeit sicherstellen zu können. Eine weitere Methode, um die Wirkung eines Angebots zu generalisieren, sind Metastudien (Eisend 2014). Die vermehrte Durchführung solcher Studien in den Arbeitsfeldern der Sozialen Arbeit in Deutschland wäre daher sehr wünschenswert.

**Resümee |** Abschließend kann festgestellt werden, dass die dargestellten Grenzen der Wirkungsanalyse keine Argumente dafür sind, sich nicht mit dem Thema Wirkungsorientierung und Wirkungsanalyse in der Sozialen Arbeit zu beschäftigen. Vielmehr ist ein differenzierter und realistischer, an der jeweiligen Praxis vor Ort orientierter Umgang damit nötig!

<sup>3</sup> Eine Übersicht über Probleme, die im Hinblick auf die interne Validität auftreten können, geben unter anderen Müller (2017) sowie Shadish et al. (2002).

Der vorgestellte Wirkungsradar als empirische Toolbox bietet genau diesen Zugang für die verschiedenen Arbeitsfelder der Sozialen Arbeit. Mit dem Aufbau eines wirkungsorientierten Monitorings und der Durchführung einer Wirkungsplausibilisierung sollte es auch dauerhaft möglich sein, die eigenen Effekte der Arbeit auch gegenüber dem Kostenträger darzustellen und die neuen gesetzlichen Anforderungen zu erfüllen.

Erste Erfahrung aus Projekten zeigen, dass die Toolbox von den Mitarbeitenden in verschiedenen Feldern der Sozialen Arbeit positiv aufgenommen wird. Vor allem die fachliche Perspektive im Rahmen der Wirkungsdebatte und die Erstellung von Wirkmodellen werden besonders positiv bewertet. Gerade bei den Wirkmodellen wird hervorgehoben, dass dadurch die Reflexion der eigenen fachlichen Praxis erfolgen kann. Dies wird als deutlicher Mehrwert angesehen. Auch wurden mit dem Schritt „Effekte identifizieren“ bereits gute Erfahrungen gemacht, da Veränderungen bei den Klientinnen und Klienten einer Maßnahme darstellbar wurden, die zum einen gegenüber Kostenträgern und Politik berichtet werden konnten, zum anderen aber auch für die eigene fachliche Entwicklung eingesetzt werden können. Hierbei zeigte sich auch, dass die Resultate sozialer Maßnahmen auf diese Weise wesentlich differenzierter dargestellt werden können und dadurch eben nicht nur an harten, quantitativen Faktoren wie beispielsweise der erfolgreichen Vermittlung in ein Arbeitsverhältnis festgemacht werden.

Aufgrund der methodischen Voraussetzungen für den kausalen Wirkungsnachweis muss aber auch festgestellt werden, dass das Instrumentarium in seiner vollen Breite nur dann eingesetzt werden kann, wenn Einrichtungen entsprechendes methodisches Know-how abrufen können. Hier nehmen praxisorientierte Forschungsinstitute und Hochschulen für angewandte Wissenschaften eine wichtige Rolle ein, da sie dieses Wissen zur Verfügung stellen und Einrichtungen in der Praxis im Rahmen von Wirkungsanalysen unterstützen können. Zudem wird der Prozess – je weiter man sich im Wirkungsradar hocharbeitet – komplexer. Nach allen bisherigen Erfahrungen ist es trotz allem lohnenswert, sich diesem Prozess und damit der Frage nach der Wirkung der eigenen Arbeit zu stellen. Mit seinen verschiedenen Stufen bietet der Wirkungsradar einen niederschweligen Einstieg für alle Träger und

Einrichtungen, die sich auf den Weg machen möchten, die Wirkung der eigenen Angebote besser in den Blick zu nehmen.

**Sebastian Ottmann** M.A. ist wissenschaftlicher Mitarbeiter am Institut für Praxisforschung und Evaluation der Evangelischen Hochschule Nürnberg. Im Rahmen dieser Tätigkeit führt er Wirkungsanalysen und Evaluationsstudien im Bereich der Sozialen Arbeit durch und beschäftigt sich seit mehreren Jahren mit den Themen Wirkungsorientierung und Wirkungsanalyse in der Sozialen Arbeit. E-Mail: sebastian.ottmann@evhn.de

**Prof. Dr. Joachim König** lehrt Allgemeine Pädagogik und Empirische Sozialforschung und ist Vizepräsident für Forschung und Entwicklung an der Evangelischen Hochschule Nürnberg. Er leitet dort das Institut für Praxisforschung und Evaluation. E-Mail: joachim.koenig@evhn.de

Dieser Beitrag wurde in einer Double-Blind Peer Review begutachtet und am 20.3.2019 zur Veröffentlichung angenommen.

375

## Literatur

- Angrist, J.D.; Pischke, J.-S.:** Mostly harmless econometrics: an empiricist's companion. Princeton 2009
- Balzer, L.:** Der Wirkungsbegriff in der Evaluation – eine besondere Herausforderung. In: Niedermaier, G. (Hrsg.): Evaluation als Herausforderung der Berufsbildung und Personalentwicklung. Linz 2012, S. 125-141
- Bartsch, S.; Beywl, W.; Niestroj, M.:** Der Programmbaum als Evaluationsinstrument. In: Giel, S.; Klockgether, K.; Mäder, S. (Hrsg.): Evaluationspraxis: Professionalisierung – Ansätze – Methoden. Münster 2016, S. 89-111
- Beach, D.; Pedersen, R.B.:** Process-tracing methods: foundations and guidelines. Ann Arbor 2013
- Beywl, W.; Niestroj, M.:** Der Programmbaum. Landmarke wirkungsorientierter Evaluation. In: Beywl, W.; Niestroj, M. (Hrsg.): Das ABC der wirkungsorientierten Evaluation: Glossar – deutsch/englisch – der wirkungsorientierten Evaluation. Köln 2009, S. 137-149
- Bibliographisches Institut:** Wirkung, die. In: <https://www.duden.de/rechtschreibung/Wirkung> (o.J., abgerufen am 27.6.2019)
- Borrmann, S.; Thiessen, B. (Hrsg.):** Wirkungen Sozialer Arbeit: Potentiale und Grenzen der Evidenzbasierung für Profession und Disziplin. Opladen u.a. 2016
- Caspari, A.:** „Rigoröse“ Wirkungsevaluation – methodische und konzeptionelle Ansätze der Wirkungsmessung in der Entwicklungszusammenarbeit. In: Zeitschrift für Evaluation 2/2009, S. 182-213

- Eisend, M.:** Metaanalyse. Sozialwissenschaftliche Forschungsmethoden, Band 8. München 2014
- EvalWiki:** Outputs (eines Programms) In: [\(veröffentlicht 2012, abgerufen am 14.6.2019\)](https://eval-wiki.org/glossar/Outputs_(eines_Programms))
- EvalWiki:** Monitoring. In: [\(veröffentlicht 2015, abgerufen am 27.6.2019\)](https://eval-wiki.org/w_glossar/index.php?title=Monitoring&oldid=1363)
- Giel, S.:** Theoriebasierte Evaluation: Konzepte und methodische Umsetzungen. Münster 2013
- Gollwitzer, M.; Jäger, R.S.:** Evaluation kompakt. 2. Auflage. Weinheim und Basel 2014
- Haunberger, S.; Baumgartner, E.:** Wirkungsevaluationen in der Sozialen Arbeit mittels Realistic Evaluation: empirische Anwendungen und methodische Herausforderungen. In: Zeitschrift für Evaluation 1/2017, S. 121-145
- Hensen, P.:** Qualitätsmanagement im Gesundheitswesen. 2. Auflage. Wiesbaden 2019
- Kolip, P.; Greif, N.:** Evaluation Programm Klasse 2000: Zusammenfassender Abschlussbericht. In: [\(veröffentlicht 2016, abgerufen am 11.7.2019\)](https://www.klasse2000.de/fileadmin/user_upload/Studie_Ernaehrung_und_Bewegung_2016-Abschlussbericht.pdf)
- Macsenare, M.:** Hilfen zur Erziehung – Praxis trifft Forschung, Forschung trifft Praxis: Was wir voneinander wissen (wollen). In: Forum Jugendhilfe 1/2018, S. 28-32
- Merton, R.K.:** Social Theory and Social Structure. New York 1968
- Müller, C.E.:** Kausale Wirkungsevaluation zwischen methodischem Anspruch und empirischer Praxis. In: Stockmann, R.; Meyer, W. (Hrsg.): Die Zukunft der Evaluation: Trends, Herausforderungen, Perspektiven. Münster 2017, S. 205-222
- Ottmann, S.; König, J.:** Was wirkt wie? – Konzeptionelle Überlegungen zur Messung und Analyse von Wirkungen in der Sozialen Arbeit. Der Wirkungsradar des Instituts für Praxisforschung und Evaluation der Evangelischen Hochschule Nürnberg. In: Forschung, Entwicklung, Transfer. Nürnberger Hochschulschriften 29/2018 (<https://doi.org/10.17883/fetschriften029>, abgerufen am 25.6.2019)
- Ottmann, S.; König, J.:** Am Anfang steht das Wirkmodell – Überlegungen und Ansätze zur Analyse von Wirkungen in der Sozialen Arbeit. In: Blätter der Wohlfahrtspflege (im Erscheinen 2019)
- Pawson, R.:** Middle Range Theory and Program Theory Evaluation: From Provenance to Practice. In: Vaessen, J.; Leeuw, F.L. (eds.): Mind the Gap. Perspectives on Policy Evaluation and the Social Sciences. New Brunswick und London 2010, pp. 171-202
- Pawson, R.; Tilley, N.:** Realistic evaluation. London und Thousand Oaks 1997
- Rauscher, O.; Mildenberger, G.; Krlev, G.:** Wie werden Wirkungen identifiziert? Das Wirkungsmodell. In: Schober, C.; Then, V. (Hrsg.): Praxishandbuch Social Return on Investment: Wirkung sozialer Investitionen messen. Stuttgart 2015, S. 41-57
- Rosenbaum, P.R.; Rubin, D.B.:** The Central Role of the Propensity Score in Observational Studies for Causal Effects. In: Biometrika 1/1983, pp. 41-55
- Schmidt, S.:** Nutzen einer wirkungsorientierten Zielsystematik. Bindeglied zwischen Evaluation und Programmsteuerung. In: Giel, S.; Klockgether, K.; Mäder, S. (Hrsg.): Evaluationspraxis: Professionalisierung – Ansätze – Methoden. Münster 2016, S. 159-175
- Schmitt, J.:** Schluss mit schwarzen Boxen. Zur Arbeit mit Kausalmechanismen in Evaluierungen. In: DEval Policy Brief 10/2018
- Schober, C.; Then, V. (Hrsg.):** Praxishandbuch Social Return on Investment: Wirkung sozialer Investitionen messen. 1. Auflage. Stuttgart 2015
- Shadish, W.R.; Cook, T.D.; Campbell, D.T.:** Experimental and Quasi-experimental Designs for Generalized Causal Inference. Belmont 2002
- Social Reporting Initiative e.V.:** SRS Social Reporting Standard: Leitfaden zur wirkungsorientierten Berichterstattung. In: [\(veröffentlicht 2014, abgerufen am 27.6.2019\)](http://www.social-reporting-standard.de/fileadmin/redaktion/downloads/SRS_Leitfaden_2014_DE.pdf)
- The Center for Theory of Change:** How Does Theory of Change Work? Theory of Change Community. In: <http://www.theoryofchange.org/what-is-theory-of-change/how-does-theory-of-change-work/> (o.J., abgerufen am 27.6.2019)
- Unz, D.; Taube, V.:** Implementierung von evidenzbasierten Programmen in der Sozialen Arbeit. Forschungsergebnisse, Herausforderungen und Desiderata. (Theorie, Forschung und Praxis der Sozialen Arbeit). In: Borrman, S.; Thiessen, B. (Hrsg.): Wirkung Sozialer Arbeit. Potentiale und Grenzen der Evidenzbasierung für Profession und Disziplin. Band 12. Opladen u.a. 2016, S. 125-140
- Vaessen, J.:** How complicated does the (Intervention) Model have to be? In: [\(veröffentlicht 2016, abgerufen am 27.6.2019\)](https://ieg.worldbankgroup.org/blog/how-complicated-does-intervention-model-have-be)
- Veerman, J.W.; van Yperen, T.A.:** Degrees of freedom and degrees of certainty: A developmental model for the establishment of evidence-based youth care. In: Evaluation and Program Planning 2/2007, pp. 212-221 (<https://doi.org/10.1016/j.evalprogplan.2007.01.011>, abgerufen am 25.6.2019)