



Forschung • Entwicklung • Transfer

Nürnberger Hochschulschriften
Nr. 53

Ottmann, Sebastian und König, Joachim (Hrsg.)

Wirkungen im Blick –
Wirkungsorientierung in der Sozialen Arbeit und
Sozialwirtschaft

Tagungsdokumentation der Transferkonferenz

2023

Forschung, Entwicklung, Transfer – Nürnberger Hochschulschriften

Die Nürnberger Hochschulschriften zu ‚Forschung, Entwicklung, Transfer‘ sind eine lose Hochschulreihe der Evangelischen Hochschule Nürnberg zur Publikation von Vorträgen, Artikeln und Projektberichten. Auch herausragende Qualifizierungsarbeiten von Studierenden können in gekürzter Fassung veröffentlicht werden. Die Online-Reihe soll dazu dienen, das Profil der Forschung und Entwicklung an der Hochschule nach innen und außen zu kommunizieren und die breite Expertise der Mitglieder der Hochschule deutlich zu machen.

Evangelische Hochschule Nürnberg
Bärenschanzstraße 4
90429 Nürnberg

www.evhn.de

Sebastian Ottmann M.A., Prof. Dr. Joachim König

Wirkungen im Blick – Wirkungsorientierung in der Sozialen Arbeit und Sozialwirtschaft.
Tagungsdokumentation der Transferkonferenz

Nürnberg, 2023

Zitiervorschlag:

Ottmann, S. & König, J. (Hrsg.) (2023): Wirkungen im Blick – Wirkungsorientierung in der Sozialen Arbeit und Sozialwirtschaft. Tagungsdokumentation der Transferkonferenz. *Forschung, Entwicklung, Transfer – Nürnberger Hochschulschriften, Nr. 53*. Nürnberg: Evangelische Hochschule Nürnberg. doi: 10.17883/fet-schriften053.

Inhaltsverzeichnis

Wirkungen im Blick, aus verschiedenen Perspektiven – ein kurzes Vorwort.....	4
Wirkungsorientierung in der Sozialen Arbeit – Herausforderung und Chance für das Arbeitsfeld	6
„Die Zahlen stehen jedoch wieder drüber“ Nichtintendierte Folgen einer Indikatoren gestützten Wirkungsorientierung in der Beratung	16
Soziale Arbeit als Investitionsobjekt? Private wirkungsorientierte Investitionen in Soziale Dienste	27
Soziale Arbeit wirkt – und es gibt Wege, dies zu zeigen. Erfahrungen aus Evaluationen	39
Evaluationen in Organisationen der Sozialen Arbeit – Bedingungen und Faktoren der Angleichung und Variation von Evaluationsaktivitäten	50
Was wirkt für wen, wie und unter welchen Bedingungen? Wirkungsmodelle, Wirkfaktoren und Wirkungen in der Bewährungshilfe im Blick.....	61
Selbstevaluation als Mittel zur Wirkungsforschung in der Sozialen Arbeit	72
Was ist Gelingen? Erkenntnisse eines Modellprojekts zur Arbeitsmarktintegration von Frauen mit Migrationshintergrund.....	83
Wissenschaftsvermittlung an HAWs – nur ein Verständnisproblem?	97
Wirkungen im Blick, aus verschiedenen Perspektiven – ein kurzes Fazit.....	108

Joachim König

Wirkungen im Blick, aus verschiedenen Perspektiven – ein kurzes Vorwort

Wirkungen im Blick – so lautete das Motto einer Transferkonferenz am Institut für Praxisforschung und Evaluation (IPE) der Evangelischen Hochschule Nürnberg im Februar 2021 zu den Ansätzen und aktuellen Herausforderungen der Wirkungsorientierung in der Sozialen Arbeit und der Sozialwirtschaft. Es entstand ein anregender und konstruktiver Dialog auf Augenhöhe zwischen Vertreter*innen der Praxis, der Verbände und der Politik einerseits sowie Akteur*innen aus Forschung, Lehre und Sozialwissenschaften. In den insgesamt elf Beiträgen, die Gegenstand dieses Tagungsbandes sind, wurde die große Bandbreite der inzwischen differenzierten empirisch-methodischen, praktisch-konzeptionellen sowie der strategischen und sozialpolitischen Debatten zu diesem Themenkreis deutlich.

Nicht nur im Zusammenhang mit der Novelle des Bundesteilhabegesetzes wurden in den letzten Monaten und Jahren Ansätze und Positionen erarbeitet, die zur Klärung der Frage beitragen sollen,

- was es eigentlich empirisch-sozialwissenschaftlich bedeutet, die Wirkung einer Praxis, einer Maßnahme, als *kausalen* Zusammenhang zu behaupten und tatsächlich *empirisch erfassen und seriös* messen zu wollen,
- und was zu tun ist, damit die Befunde aus solchen Messungen in der Praxis dann auch eine möglichst Gewinn bringende Anwendung finden – zur Entwicklung von Strategien, Konzepten und neuen Methoden oder in der Organisationsentwicklung.

Unter anderem wurde dazu am IPE ein empirischer Methodenkoffer, genannt Wirkungsradar, entwickelt und in verschiedenen Pilotprojekten – geleitet von einem zweifachen Anspruch – umgesetzt:

- Zum einen das Instrumentarium möglichst praxisingängig, also mit einem überschaubaren Aufwand im Alltagsgeschäft, in den Einrichtungen umsetzbar und einsetzbar zu machen,
- und zum anderen die fachlich-konzeptionelle Perspektive bei der Modellbildung und Operationalisierung immer zu berücksichtigen, also durch die Beteiligung der Fachkräfte und gegebenenfalls auch der Klient*innen die Validität der Befunde zu optimieren.

Vor dem Hintergrund dieser und aller weiteren inzwischen gesammelten Erfahrungen war es die Aufgabe der Konferenz, sowohl aus der Perspektive der Praxis als auch mit der sozialwissenschaftlichen Brille in einen Austausch und in eine differenzierte Debatte zu kommen – vor dem Hintergrund von sieben zentralen Fragen:

1. Wie lässt sich der Gegenstand, also die Wirkung einer Maßnahme, am besten von den weiteren in dem Zusammenhang üblichen Begriffen wie Erfolg, Effektivität und Effizienz, Output und Outcome ab- und eingrenzen?
2. Welche Modelle können Wirkungen am besten abbilden? Gibt es sinnvolle generelle Indikatoren-Sets für bestimmte Arbeitsfelder, die in einem weiteren Schritt in diesen Feldern spezifizierbar sind?
3. Welche empirischen Designs und welche quantitativen und qualitativen Methoden sind dafür am besten geeignet?
4. Welche Erfolgsprädiktoren für die Implementierung einer ‚Organisationskultur der Wirkungsorientierung‘ in sozialen Diensten, hinein in die Köpfe und Herzen der Fachkräfte,

lassen sich formulieren, etwa mit Blick auf Rahmenbedingungen, auf das Commitment und auf die notwendige professionelle Haltung bei den Fachkräften?

5. Woran kann die Praxisgängigkeit von Erhebungsverfahren festgemacht werden, z.B. bezogen auf den notwendigen Aufwand im Alltagsgeschäft oder die Fehleranfälligkeit und die möglichen Verzerrungseffekte bei den Messungen?
6. Was muss beim Aufbau eines fachkräftefreundlichen Datenmanagements in den Einrichtungen beachtet werden? Wie sieht deswegen ein sinnvolles Monitoring aus? Wie lassen sich z.B. Dashboards aufbauen, die vor Ort für die Akteure leicht handhabbar sind? Welche Datenkompetenz braucht es dazu bei den Fachkräften? Welcher Qualifizierungsbedarf leitet sich daraus für die Curricula der Hochschulen und für die Fort- und Weiterbildung ab?
7. Welche Strategien gibt es, die Befunde aus Wirkungsanalysen möglichst adressaten- und nutzerfreundlich in die Qualitäts- und Organisationsentwicklung von Einrichtungen und auch in die verbandlichen und politischen Debatten zu transferieren?

Die folgenden Beiträge können dazu weiterführende Perspektiven eröffnen und einen Beitrag leisten, sowohl die Debatte der scientific community zu befeuern als auch das Ringen der Leistungserbringer sowie der Leistungsträger in den verschiedenen Feldern der Sozialwirtschaft, um gute Lösungen in konstruktiver Weise voranzubringen. Am IPE freuen wir uns darauf, in diesem Zusammenhang auch in der kommenden Zeit den einen oder anderen hilfreichen Beitrag zur Weiterentwicklung leisten zu können.

Prof. Dr. Joachim König ist Professor für allgemeine Pädagogik und empirische Sozialforschung an der Evangelischen Hochschule Nürnberg. Er leitet das Institut für Praxisforschung und Evaluation und ist Vizepräsident für Forschung.

Sebastian Ottmann

Wirkungsorientierung in der Sozialen Arbeit – Herausforderung und Chance für das Arbeitsfeld

In den letzten Jahren gibt es auf der einen Seite wieder eine stärkere Diskussion über die Wirkung und die Wirksamkeit von Angeboten der Sozialen Arbeit und wie diese empirisch erfasst werden können (Borrmann & Thiessen, 2016; Eppler, Miethé & Schneider, 2011; Otto, Polutta & Ziegler, 2007; Schober & Then, 2015; Schulze-Krüdener, 2017), zum anderen wird aber auch das Thema von Seiten des Kostenträgers verstärkt verfolgt. So wurde mit Einführung des Bundesteilhabegesetzes (BTHG) innerhalb der Eingliederungshilfe die Wirkung, Wirksamkeit und eine Wirkungskontrolle explizit im Gesetzaufgenommen. Durch die Corona-Krise in den letzten Jahren und die enormen finanziellen Ausgaben ist leider auch anzunehmen, dass zukünftig, bei weniger zur Verfügung stehenden finanziellen Mitteln bei Kommunen und Ländern, auch die Finanzierung von sozialen Angeboten vermehrt kritisch hinterfragt wird und damit der Wirkung des Angebots eine besondere Bedeutung zukommt.

In der Praxis zeigt sich allerdings, dass es rund um den Themenkomplex Wirkung und Wirksamkeit noch viele offene Fragen gibt. Angefangen von der Frage, wie die Wirkung und Wirksamkeit der eigenen Arbeit empirisch nachgewiesen werden kann, bis zur Tatsache, dass mit dem Thema auch Ängste formuliert werden, dass eine stärkere Fokussierung auf Wirkungen auch negative Auswirkungen auf die eigene Arbeit haben kann, wenn eine solche nicht nachgewiesen werden kann, z. B. durch finanzielle Kürzungen. In diesem Beitrag sollen daher Herausforderungen, aber auch Chancen, die sich mit dem Konzept der Wirkungsorientierung in den Arbeitsfeldern der Sozialen Arbeit ergeben, näher betrachtet werden.

Bei der Beschäftigung mit dem Konzept der Wirkungsorientierung empfiehlt es sich, in einem ersten Schritt die Definition des Begriffs Wirkung sich zu vergegenwärtigen. Im Duden wird Wirkung als eine Veränderung, die durch eine verursachende Kraft bewirkt wurde, definiert (vgl. Bibliographisches Institut, o. J.). Zentral in dieser Definition ist der kausale Mechanismus, der beschrieben wird. Für die Soziale Arbeit bedeutet dies, dass wenn von Wirkung gesprochen wird, man nicht nur die Veränderung bei der Zielgruppe in den Blick nehmen muss, sondern auch mit einer sehr hohen Wahrscheinlichkeit davon ausgehen muss, dass diese Veränderung durch das durchgeführte Angebot oder die durchgeführte Maßnahme entstanden ist und nicht durch andere Faktoren. Dieser kausale Mechanismus stellt sehr hohe Anforderungen an den Wirkungsnachweis dar, wie wir in diesem Beitrag noch sehen werden. An dieser Stelle sei noch angemerkt, dass nicht immer Veränderungen bei der Zielgruppe auftreten müssen, um eine Wirkung darstellen zu können. Auch Stabilisierungen eines aktuellen Status kann eine Wirkung darstellen (vgl. Balzer & Beywl, 2015, S. 192), allerdings ist auch hier der Nachweis des kausalen Mechanismus nötig.

Um differenziert mit dem Begriff Wirkung umzugehen, erscheint daher eine begriffliche Unterscheidung sinnvoll, je nachdem was man empirisch einlösen kann. In der Arbeit am Institut für Praxisforschung und Evaluation hat sich nachfolgende Unterscheidung als praxisgänglich erwiesen (vgl. Ottmann & König, 2022):

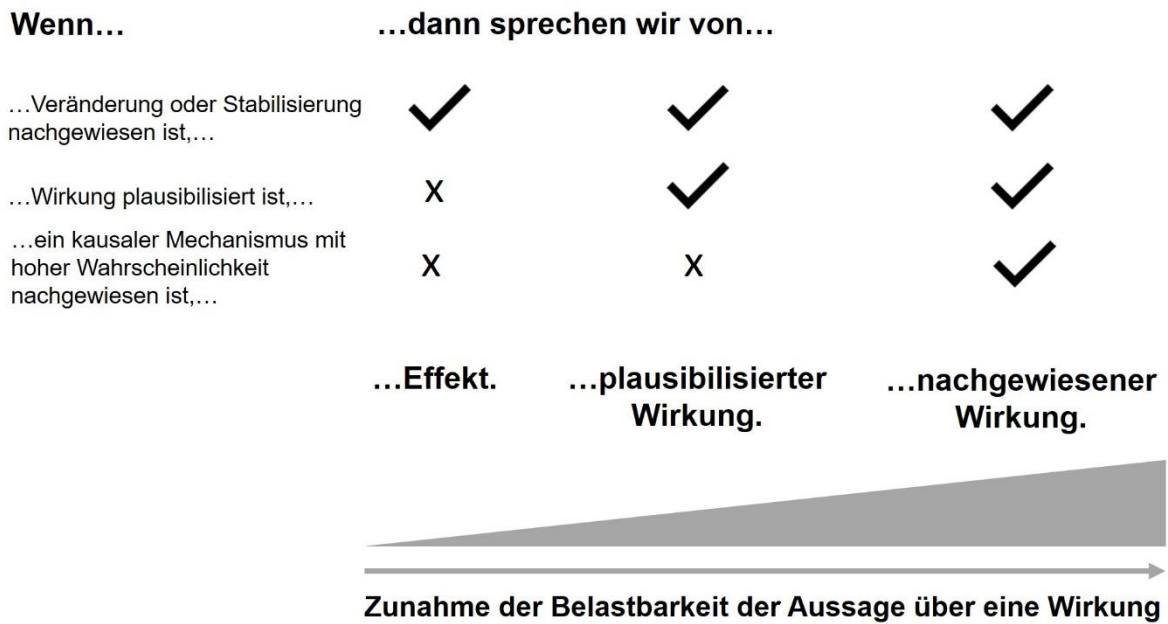


Abbildung 1: Vorschlag für einen differenzierten Umgang mit dem Begriff Wirkung

Die Begrifflichkeit die verwendet wird hängt davon ab, was empirisch erhoben und nachgewiesen werden kann. Wurde in einem ersten Schritt eine empirische Erhebung durchgeführt, die die Veränderungen oder Stabilisierungen auf der Ebene der Zielgruppe valide nachweisen kann, wird von einem Effekt gesprochen.

Um von einer nachgewiesenen Wirkung zu sprechen, muss der kausale Mechanismus mit hoher Wahrscheinlichkeit nachgewiesen werden. Hierzu wird mindestens ein quasi-experimentelles Vergleichsgruppendesign benötigt (siehe hierzu auch den dazugehörigen Abschnitt in diesem Beitrag). In der Praxis der Sozialen Arbeit ist dies oft nur sehr schwer umzusetzen. Wenn dies der Fall ist, sollte aber neben der Erfassung der Veränderung und Stabilisierungen eine sog. Wirkungsplausibilisierung durchgeführt werden. Balzer (2012) führte diesen Begriff als Abgrenzung zum empirischen Wirkungsnachweis ein, wenn keine kausalen Ursachen-Wirkungs-Zusammenhänge vorliegen oder nachgewiesen werden können. Bei dieser Methode geht es darum, die Wirkung einer Maßnahme zu plausibilisieren, zum Beispiel durch die Interpretation der erhobenen Daten mit Fachkräften oder die Erhebung von weiteren empirischen Daten. Wurde eine solche Plausibilisierung durchgeführt, kann man von einer plausibilisierten Wirkung sprechen. Nach den bisherigen Erfahrungen am Institut für Praxisforschung und Evaluation der Evangelischen Hochschule Nürnberg stellt die plausibilisierte Wirkung einen guten und gangbaren Weg in der Praxis dar, die Wirkung und Wirksamkeit der eigenen Arbeit nachzuweisen.

Chance: Perspektiven von Wirkungsorientierung

In Projekten im Bereich der Wirkungsorientierung und der Wirkungsanalyse zeigt sich, dass die Perspektive, mit der ein solches Projekt angegangen wird, zentral ist, vor allem auch für das Commitment der beteiligten Fachkräfte. Der Zugang zu diesem Thema kann aus unterschiedlichen Perspektiven erfolgen, wie die nachfolgende Abbildung aufzeigt (vgl. Ottmann & König, 2023):



Abbildung 2: Perspektiven von Wirkungsorientierung

Oft wird das Thema Wirkungsorientierung aus einer legitimatorischen Perspektive angegangen, weil beispielsweise der Kostenträger den Nachweis der Wirkung der finanzierten Angebote einfordert. Gerade diese Perspektive ist es aber auch, die Ängste unter den Fachkräften auslöst, weil oft nicht geklärt ist, wie mit dem Fall umgegangen wird, dass eine gewünschte Wirkung nicht eintritt. Daher sollte die Beschäftigung mit der Wirkung der eigenen Angebote immer zuerst aus einer fachlichen-konzeptionellen Perspektive erfolgen. Denn in dieser Perspektive liegt die Chance, mit den Erkenntnissen aus dem Prozess der Wirkungsorientierung und der Durchführung von Wirkungsanalyse Anhaltspunkte für die Weiterentwicklung des eigenen Handelns zu erhalten. Auch kann mit der Zeit ein Wissenskorpus aufgebaut werden, mit dem die Fachkräfte dann die besten Angebote oder Maßnahmen für die Klient*innen identifizieren können. Dies sorgt nicht zuletzt auch dafür, dass durch die Beschäftigung mit der eigenen Wirkung eine weitere Professionalisierung des Arbeitsbereiches erfolgt (vgl. Lehmann, 2020). In der Praxis zeigt sich auch, dass dieser Zugang auch positiv von den Fachkräften gesehen wird, weil durch diesen eine Reflektion der eigenen Arbeit erfolgen kann. Daher sollte man das Konzept der Wirkungsorientierung und der Wirkungsanalyse als Chance für die eigene Arbeit ansehen.

Chance: Stufenweises Vorgehen

Die Begriffe Wirkung und Wirkungsorientierung werden sehr oft und sehr schnell synonym mit dem Begriff Wirkungsanalyse verwendet. Dadurch wird davon ausgegangen, dass bei der Beschäftigung mit diesem Thema immer auch eine groß angelegte empirische Erhebung erfolgen muss. Dass dies nicht so sein muss, zeigt der Wirkungsradar, der am Institut für Praxisforschung und Evaluation der Evangelischen Hochschule Nürnberg entwickelt wurde und ein stufenweises Vorgehen vorschlägt (vgl. Ottmann & König, 2018, 2019a).

Der IPE Wirkungsradar



Abbildung 3: Der Wirkungsradar des Instituts für Praxisforschung und Evaluation

Zentral im Wirkungsradar ist die Verbindung von sogenannten Black-Box- und White-Box-Verfahren. Es soll also nicht nur die Frage, *ob* eine Maßnahme oder ein Angebot wirkt beantwortet werden, sondern auch die Frage nach dem *wie*. Zum Einstieg wird empfohlen, in einem Wirkungsworkshop gemeinsam innerhalb einer Einrichtung oder eines Trägers das Verständnis zum Thema Wirkung und Wirkungsorientierung zu klären. Danach können Wirkmodelle erstellt und entwickelt werden. In einem Wirkmodell wird theoretisch festgehalten, mit welchen Wirkungen man bei einem Angebot oder einer Maßnahme rechnet, wie diese entstehen und in welchem Kontext diese eingebunden sind. In diesem zweiten Schritt werden noch keine empirischen Daten erhoben, sondern es erfolgt eine ‚theoretische‘ Auseinandersetzung mit den gewünschten Wirkungen. Gerade diese Entwicklung ist für Einrichtungen und Träger mit einem vertretbaren Aufwand möglich. Danach können darauf aufbauend empirische Untersuchungen im Rahmen von Wirkungsanalysen durchgeführt werden. Wichtig hierbei ist aber die Unterscheidung zwischen der Identifizierung von Effekten und dem eigentlichen kausalen Nachweis von Wirkungen, wie er schon bei der Begriffsbestimmung am Anfang des Beitrags dargestellt wurde. In einem vorletzten Schritt können dann Wirkmechanismen näher analysiert werden und wenn gewünscht die Effizienz des Angebotes dargestellt werden.

In der Praxis hat sich gezeigt, dass die Entwicklung von Wirkmodellen schon positive Auswirkungen auf die eigene Arbeit hat und auch eine Möglichkeit der Darstellung der eigenen Arbeit bietet (vgl. Ottmann & König, 2019b). Die Erstellung eines solchen Modells kann zum einen die eigene Arbeit reflektieren, zum anderen wird aber auch durch die Darstellung deutlich, wie komplex oft soziale Interventionen sind. Es kann daher deutlich gemacht werden, dass zum

Erreichen einer bestimmten Hauptwirkung (z. B. den Übergang in den ersten Arbeitsmarkt) eine Reihe von Teilwirkungen entstehen (z. B. ein Kompetenzzuwachs) und dieser auch entsteht, wenn die Hauptwirkung am Ende nicht erreicht wird.

Durch das im Wirkungsradar dargestellte stufenweise Vorgehen, kann der Einstieg in die Thematik in den Einrichtungen und bei Trägern einfach erfolgen und die einzelnen Schritte können je nach vorhandenen Ressourcen durchlaufen werden.

Herausforderung: methodische Anforderungen an den Wirkungsnachweis

Möchte man eine Wirkung empirisch nachweisen, muss ein Nachweis des kausalen Mechanismus erfolgen. Dieser Nachweis hat allerdings hohe Anforderungen an das methodische Design. Dies ergibt sich daraus, da empirisch nachgewiesen werden muss, dass lediglich das Angebot oder die Maßnahme für die Veränderungen oder Stabilisierungen bei den Klient*innen verantwortlich ist und keine anderen Faktoren, wie beispielsweise das familiäre Umfeld. Gerade bei Angebot und Dienstleistungen der Sozialen Arbeit ist dieser Nachweis sehr herausfordernd, da diese oft in einer sehr komplexen sozialen Wirklichkeit eingebettet sind mit vielen weiteren Faktoren, die einen positiven Ausgang bedingen können.

Um einen kausalen Mechanismus empirisch nachweisen zu können, wird das kontrafaktische Modell der Kausalität zugrunde gelegt. Dieses besagt, dass zur Erfassung des kausalen Effekts bei einer Person zwei Messungen an zwei Zeitpunkten stattfinden müssen: einmal, wenn die Maßnahme erfolgt ist und einmal, wenn die Person nicht an der Maßnahme teilnimmt. Der kausale Effekt der Maßnahme kann dann durch die Differenz der beiden Messwerte bestimmt werden (vgl. Morgan & Winship, 2015, S. 4). Beim Lesen dieser Annahme merkt man aber recht schnell, dass dies in der Praxis nicht möglich ist. Entweder eine Person nimmt an einer Maßnahme teil oder nicht, von daher können an einer Person nicht beide Zustände gemessen werden. Diese Problematik wird auch als das fundamentale Problem der kausalen Inferenz beschrieben (vgl. Holland, 1986). Um dieses Problem zu lösen, werden nicht Einzelpersonen betrachtet sondern Gruppen, einmal eine Gruppe die an der Maßnahme teilgenommen hat und einmal eine Gruppe die nicht an der Maßnahme teilgenommen hat (vgl. Morgan & Winship, 2015, S. 46 ff.). Daher besteht die Herausforderung bei einem empirischen Wirkungsnachweis in der Sozialen Arbeit darin, nicht nur die Veränderungen oder Stabilisierungen bei den Klient*innen des Angebotes zu erfassen, sondern auch bei einer Vergleichsgruppe. Die Personen dieser Vergleichsgruppe sollten den Klient*innen möglichst ähnlich sein, allerdings nicht an der Maßnahme teilnehmen können. In der empirischen Sozialforschung wird oft das Experiment als „Goldstandard“ bezeichnet, um solche Gruppen zu bilden. Hierbei entscheidet nur der Zufall, ob eine Person an der Maßnahme teilnehmen darf oder an der Kontrollgruppe. Dieses Verfahren ist in den Arbeitsfeldern der Sozialen Arbeit aus ethischen Gründen nicht durchführbar, da Personen, die einen Hilfebedarf haben, nicht aus forschungsmethodischen Gründen eine Maßnahme oder ein Angebot vorenthalten werden darf. Eine weitere Möglichkeit besteht in der Durchführung von einem quasi-experimentellen Design. Hierbei ist die Vergleichsgruppe eine schon bestehende Gruppe bzw. eine Gruppe die mit Personen gebildet wird, die nicht an der Maßnahme teilnehmen. Mit statistischen Verfahren, wie dem Propensity Score Matching, kann eine solche Vergleichsgruppe auch so zusammengestellt werden, dass diese der Maßnahmengruppe recht ähnlich ist. Bei einer guten Datenlage und einer guten Durchführung dieser Methode können ähnlich valide Ergebnisse wie mit

Experimentalstudien erzielt werden (vgl. Caspari, 2009). Allerdings stellt auch dieses Vorgehen die Herausforderung dar, an die Personen zu kommen, die einer Vergleichsgruppe angehören. Es stellt sich die Frage, wie Personen erreicht werden können, die einen bestimmten Hilfebedarf haben, aber bisher keine Angebote oder Dienstleistungen der Sozialen Arbeit in Anspruch nehmen. Daher ist es in vielen Arbeitsfeldern schwierig, solche Vergleichsgruppen zu bilden¹ und es empfiehlt sich, mit der schon dargestellten Methode der Wirkungsplausibilisierung zu arbeiten.

Herausforderung: einheitliche Indikatorensets

Eine Frage, die sehr oft in der Praxis gestellt wird, ist die nach einheitlichen Indikatorensets für die Wirkungsanalyse. Hier muss man allerdings feststellen, dass aktuell in den Arbeitsbereichen die eingesetzten Indikatoren eine Aushandlung zwischen den Beteiligten, also Leistungserbringer und Kostenträger, bedarf und es keine einheitlichen Indikatoren gibt die ‚sofort‘ herangezogen werden können. Die Entwicklung solcher einheitlichen Indikatorensets wird eine der Aufgaben der wissenschaftlichen Forschung in den nächsten Jahren sein. Dabei stellen sich aber zwei Fragen: Wie einheitlich können solche Indikatorensets sein? Hier muss zwischen den individuellen Bedürfnissen von Einrichtungen auf der einen Seite und der Vergleichbarkeit der Ergebnisse auf der anderen Seite abgewogen werden.

Welche Bereiche und Wirkungen werden mit diesen Indikatoren in den Blick genommen?

Gerade die letzte Frage bedarf einer umfassenderen Diskussion. Aktuell zeigt sich, dass bei der Frage nach der Wirksamkeit oft Indikatoren gewählt werden, die einfach zu bestimmen sind, z. B. die Übergangsquote in den ersten Arbeitsmarkt. Für die umfängliche Bewertung der Wirksamkeit von sozialen Angeboten und Dienstleistungen greifen diese aber oft zu kurz. Wird diese Übergangsquote nicht erfüllt, bedeutet es nämlich nicht automatisch, dass das Angebot keine Wirkung auf Ebene der Klient*innen erzielt hat. Viel mehr wird hier ein differenzierter Mix an fachlichen Indikatoren benötigt, die beispielsweise auch Kompetenzzuwächse darstellen, die entstanden sind.

Bei der Frage nach einheitlichen Indikatoren könnte ein zukünftiges Vorgehen in einem ersten Schritt bei der Entwicklung von sog. ‚Ankerwirkmodellen‘ liegen. Diese ‚Ankerwirkmodelle‘ werden für einzelne Arbeitsbereiche oder Angebote entwickelt und haben den Anspruch, eine Allgemeingültigkeit darzustellen in dem Sinne, dass diese in verschiedenen Einrichtungen eingesetzt werden können, die ein entsprechendes Angebot anbieten. Der Vorteil solcher ‚Ankerwirkmodelle‘ wäre, dass nicht jede Einrichtung bzw. jeder Träger ein komplett neues Wirkmodell entwickeln müsste, sondern auf dieses ‚Ankerwirkmodell‘ aufbauen kann und lediglich entsprechende Ergänzungen oder Erweiterungen vornehmen muss. Ähnliche Überlegungen und auch konkrete Vorschläge finden sich in der Literatur bei Funnell und Rogers (2011, S. 351 ff.) mit sogenannten Archetypen für bestimmte Interventionen und bei Wachsmut, Brinkmann und Hense (2019, S. 298 ff.) als Wirkungsmodelltypen für die Non-Profit-Programmarbeit. Eine erste Entwicklung von Ankerwirkmodellen in den Arbeitsfeldern der Sozialen Arbeit liegen im Bereich der Eingliederungshilfe für das Angebot des Berufsbildungsbereiches und der teilstationären

¹ Eine Möglichkeit kann ein Wartekontrollgruppendesign sein. In diesem wird das Angebot bei dem Wirkung untersucht wird nicht gleichzeitig in allen Gruppen eingeführt, sondern nur bei einem Teil der Gruppen. Die anderen Gruppen erhalten dann das neue Angebot erst nach einer festgelegten Zeit, beispielsweise nach einem Jahr.

Tagesbetreuung für erwachsene Menschen mit Behinderung nach dem Erwerbsleben vor (vgl. Ottmann, König & Gander, 2021). Ein weiterer Schritt könnte dann sein, aufbauend auf diesen Ankerwirkmodellen Indikatoren zu entwickeln, die dann auch in den Einrichtungen eingesetzt werden. Prinzipiell wäre es sinnvoll, diese Indikatoren dann auch in ein IT-System zu überführen, das für die Einrichtungen einen leichten Zugang zur Datenerhebung und Datenauswertung ermöglicht. Um solche Indikatoren zu entwickeln ist aber noch weitere Forschungsarbeit nötig, gerade auch im Hinblick auf die testtheoretische Absicherung (vgl. Eid & Schmidt, 2014) der entwickelten Indikatoren.

Herausforderung: Meta-Studie

Bei der Diskussion über die Erfassung von Wirkungen der Sozialen Arbeit wird sehr oft über die empirische Erhebung auf Ebene der Einrichtung oder des Trägers diskutiert. Eine Herausforderung in den kommenden Jahren wird hier sein, den Blick von Einrichtungsebene auf die Ebene des Arbeitsfeldes zu lenken. So wäre es wünschenswert, wenn zukünftig vermehrt Meta-Studien (vgl. Eisend, 2014) in den verschiedenen Arbeitsfeldern der Sozialen Arbeit und Sozialwirtschaft durchgeführt werden. Hierbei werden Studien innerhalb eines Arbeitsfeldes gesammelt und miteinander verglichen, um zu allgemeingültigen Aussagen über die Wirksamkeit in einem Arbeitsbereich zu kommen.

Wünschenswert wäre natürlich, wenn die Ergebnisse solcher Studien auch in allgemeinen zugänglichen Datenbanken veröffentlicht werden. Hierdurch kann ein umfangreicher Wissenskorpus für die Soziale Arbeit aufgebaut werden, auf den dann auch wieder Fachkräfte in ihrer täglichen Arbeit zurückgreifen können. Ein gelungenes Beispiel für eine solche Datenbank ist das Projekt „WESPE – Wissenschaftliche Erkenntnisse zu Sicherheits- und Präventionsmaßnahmen durch Evaluation“, das Informationen über die Wirksamkeit kriminalpräventiver Maßnahmen in Deutschland in einer Datenbank aufbereitet (vgl. Armbrorst & Walsh, 2019).

In diesem Zusammenhang stellt sich aber auch die Frage, welche Institution diese Meta-Studien durchführt und wie die Finanzierung von solchen Projekten erfolgt. Hier wäre es wünschenswert, wenn eine unabhängige Stelle geschaffen bzw. beauftragt wird, die solche Studien durchführt, ähnlich wie dies auch im Bereich der Medizin der Fall ist (vgl. Boecker & Weber, 2021, S. 57).

Herausforderung: Datenkompetenz

Um die Ergebnisse von Wirkungsanalyse und allgemein von empirischen Studien und Evaluationen in die tägliche Arbeit zu integrieren, wird es zukünftig wichtiger werden, dass die Fach- und Führungskräfte in der Sozialen Arbeit und Sozialwirtschaft über eine umfassende Datenkompetenz verfügen. Um diese zu stärken, wird es nötig sein, dass es Entwicklungen bei den Akteur*innen, die Evaluationen durchführen gibt, aber auch innerhalb der Einrichtungen und Träger.

Um die Arbeit mit den Ergebnissen innerhalb der Einrichtungen zu fördern, wird es zukünftig bedeutender sein, die Ergebnisse adressatenbezogen aufzubereiten. Hier zeigt sich, dass mit den neuen technischen Möglichkeiten auch eine bessere Darstellung von Ergebnissen möglich ist. Zukünftig erscheint es daher sinnvoll, die Ergebnisse auch in Form eines interaktiven Dashboards zur Verfügung zu stellen (vgl. Ottmann, 2021). Hierbei haben die Fachkräfte die Möglichkeit, selbst sich die Analysen zusammenstellen und bei aufkommenden Fragestellungen selbständig

Auswertungen, in einem zuvor definierten Umfang, vorzunehmen. So können in einem Dashboard beispielsweise die Ergebnisse von verschiedenen Erhebungsjahren berücksichtigt werden, die dann von den Fachkräften individuell verglichen und analysiert werden können. Ein weiterer Vorteil eines solchen Dashboards ist, dass eine Datenaktualisierung automatisiert durchgeführt werden kann, mit Anbindung an einen Online-Fragebogen sogar immer mit dem aktuell erhobenen Datenstand.

Um Ergebnisse aber gut interpretieren zu können und auch Ableitungen für das eigene fachliche Handeln zu ermöglichen, muss auch die Datenkompetenz in den Einrichtungen gestärkt werden. Hierzu wird es zukünftig verstärkt Schulungen geben müssen und auch schon im Rahmen der Ausbildung und im Studium von zukünftigen Fachkräften verstärkt auf dieses Thema Bezug genommen werden. Einen guten Handlungsleitfaden stellt hier das veröffentlichte Framework für Data Literacy dar (vgl. Schüller, Busch & Hindinger, 2019).

Herausforderung: dauerhafte Implementierung

Die zuvor ausgeführte Herausforderung der Datenkompetenz stellt auch eine Voraussetzung für die dauerhafte Implementierung von Wirkungsorientierung und Wirkungsanalyse in sozialen Einrichtungen und bei deren Trägern dar. Denn dies sollte das Ziel sein, nicht die einmalige Durchführung einer Wirkungsanalyse, sondern die dauerhafte Implementierung von Wirkungsorientierung als eine Grundlage für die fachliche Arbeit. Hierbei ist es vor allem wichtig, dass es einen regelmäßigen Kreislauf von Datenerhebung und Datenanalyse und der Interpretation dieser Daten und die Ableitung von möglichen Weiterentwicklungen für das Angebot gibt. Dies kann u. a. mit einem wirkungsorientierten Monitoring (vgl. Ottmann & König, 2018, 2019a) geleistet werden, mit dem regelmäßig Daten über die Veränderungen und Stabilisierungen der Klient*innen erhoben werden. Hierbei wird sich auch die Frage stellen, wie sich ein solches System im Rahmen der bestehenden IT abbilden lässt und als Fachcontrolling dann auch mit anderen Steuerungsinstrumenten in den Einrichtungen oder bei dem Träger verbunden werden kann, z. B. mit dem finanziellen Controlling. Die Ergebnisse eines solchen wirkungsorientierten Monitorings sollten dann in regelmäßigen Abständen interpretiert werden, einmal im Hinblick auf eine mögliche Wirkungsplausibilisierung, aber auch unter der Fragestellung, ob alle Teilnehmenden an einem Angebot oder einer Maßnahme im gleichen Maße von deren Wirkung profitieren. Sollte dies nicht der Fall sein, lohnt sich ein genauerer Blick, ob es spezielle Methoden oder Konzepte für Teilzielgruppen braucht und damit das Angebot weiterentwickelt wird.

Wichtig für diese dauerhafte Implementierung von Wirkungsorientierung ist aber auch, dass Methoden zur Verfügung stehen, die diesen Prozess für die Fachkräfte vor Ort vereinfachen. Hier liegt eine weitere Herausforderung, in den nächsten Jahren solche Methoden zu entwickeln und zu testen. Hierbei stellt sich zum einen die Frage nach dem richtigen Format (z. B. Ganztagesveranstaltung oder Teil einer Teambesprechung), aber auch über das methodische Vorgehen, um dann Rückschlüsse für die eigene fachliche Arbeit herausziehen zu können. Auch wird zukünftig sicherlich die Frage gestellt werden müssen, wie die zusätzlich anfallenden Arbeiten und benötigten Ressourcen finanziert werden, die bei einer stärkeren Vorgabe im Bereich der Wirkungsorientierung von der Politik und dem Kostenträger anfallen.

Fazit: sich den Herausforderungen annehmen und die Chancen nutzen

In diesem Beitrag wurden Chancen und Herausforderungen betrachtet, die sich mit einer verstärkten Auseinandersetzung mit dem Konzept der Wirkungsorientierung innerhalb der Sozialen Arbeit ergeben. Auch wenn viele Herausforderungen benannt wurden, sollte dies kein Grund sein, sich nicht mit dem Konzept der Wirkungsorientierung zu beschäftigen. Vielmehr sollte man sich den Herausforderungen stellen, da im Konzept der Wirkungsorientierung eine große Chance für die fachliche Arbeit in der Sozialen Arbeit und Sozialwirtschaft liegt. Diesen Mehrwert sollte man immer im Blick haben und bei allen Aktivitäten in diesem Bereich zum Ziel haben. Denn dadurch kann das Konzept der Wirkungsorientierung und die Frage nach der Wirkung nicht nur zur Legitimation der eigenen Angebote dienen, sondern auch die eigene fachliche Arbeit weiterbringen und weiterentwickeln.

Sebastian Ottmann M.A. arbeitet als wissenschaftlicher Mitarbeiter am Institut für Praxisforschung und Evaluation der Evangelischen Hochschule Nürnberg und leitet dort das Kompetenzzentrum Wirkungsorientierung in der Sozialen Arbeit.

Literatur

- Armborst, A. & Walsh, M. (2019). WESPE – Wissenstransfer für Ergebnisse aus Evaluationsstudien. *Zeitschrift für Evaluation*, 18(1), 140–149.
- Balzer, L. (2012). Der Wirkungsbegriff in der Evaluation – eine besondere Herausforderung. In G. Niedermair (Hrsg.), *Evaluation als Herausforderung der Berufsbildung und Personalentwicklung* (1. Auflage, S. 125–141). Linz: Trauner.
- Balzer, L. & Beywl, W. (2015). *evaluiert: Planungsbuch für Evaluationen im Bildungsbereich* (1. Auflage.). Bern: hep verlag ag.
- Bibliographisches Institut. (o. J.). Wirkung, die. Zugriff am 13.4.2021. Verfügbar unter: <https://www.duden.de/rechtschreibung/Wirkung>
- Boecker, M. & Weber, M. (2021). *Wie lässt sich die Wirksamkeit von Eingliederungshilfe messen? sozialwissenschaftliche Anregungen* (Soziale Arbeit kontrovers). Berlin: Verlag des Deutschen Vereins für öffentliche und private Fürsorge e.V.
- Borrmann, S. & Thiessen, B. (Hrsg.). (2016). *Wirkungen Sozialer Arbeit: Potentiale und Grenzen der Evidenzbasierung für Profession und Disziplin* (1. Auflage). Opladen, Berlin & Toronto: Verlag Barbara Budrich.
- Caspari, A. (2009). „Rigore“ Wirkungsevaluation – methodische und konzeptionelle Ansätze der Wirkungsmessung in der Entwicklungszusammenarbeit. *Zeitschrift für Evaluation*, 8(2), 182–213.
- Eid, M. & Schmidt, K. (2014). *Testtheorie und Testkonstruktion* (Bachelorstudium Psychologie). Göttingen: Hogrefe.
- Eisend, M. (2014). *Metaanalyse* (Sozialwissenschaftliche Forschungsmethoden) (1. Auflage., Band 8). München Mering: Rainer Hampp Verlag.
- Eppler, N., Miethe, I. & Schneider, A. (2011). *Qualitative und quantitative Wirkungsforschung: Ansätze, Beispiele, Perspektiven* (1. Auflage). Opladen: Verlag Barbara Budrich.
- Funnell, S. & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. San Francisco, CA: Jossey-Bass.

- Holland, P. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396), 954–960.
- Lehmann, R. (2020). Evidenz als Basis der Profession. *Soziale Arbeit*, 69(6), 217–223.
- Morgan, S. L. & Winship, C. (2015). *Counterfactuals and causal inference: methods and principles for social research* (Analytical methods for social research) (2. Auflage). New York, NY: Cambridge University Press.
- Ottmann, S. (2021). Dashboard-Lösungen als Instrument zur Beteiligung von Fachkräften in Evaluationen und Wirkungsanalysen. *Zeitschrift für Evaluation*, 20(1), 188–196. <https://doi.org/10.31244/zfe.2021.01.09>
- Ottmann, S. & König, J. (2018). 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. *Forschung, Entwicklung, Transfer – Nürnberger Hochschulschriften*, 29. <https://doi.org/10.17883/fet-schriften029>
- Ottmann, S. & König, J. (2019a). Wirkungsanalyse in der Sozialen Arbeit. Differenzierung ist nötig. *Soziale Arbeit*, 68(10), 368–376.
- Ottmann, S. & König, J. (2019b). Am Anfang steht das Wirkmodell. Überlegungen und Ansätze zur Analyse von Wirkungen in der Sozialen Arbeit. *Blätter der Wohlfahrtspflege*, 166(2), 67–70. <https://doi.org/10.5771/0340-8574-2019-2-67>
- Ottmann, S. & König, J. (2023). Wirkungsorientierung in der Sozialen Arbeit. Eine Einführung für Studium und Praxis (Grundwissen Soziale Arbeit) (1. Auflage, Band 45). Stuttgart: Kohlhammer Verlag (im Druck).
- Ottmann, S., König, J. & Gander, C. (2021). Wirkungsmodelle in der Eingliederungshilfe. *Zeitschrift für Evaluation*, 20(2), 317–331. <https://doi.org/10.31244/zfe.2021.02.04>.
- Otto, H.-U., Polutta, A. & Ziegler, H. (2007). *What Works – Welches Wissen braucht die Soziale Arbeit?: Zum Konzept evidenzbasierter Praxis* (1. Auflage). Opladen: Budrich.
- Schober, C. & Then, V. (Hrsg.). (2015). *Praxishandbuch Social Return on Investment: Wirkung sozialer Investitionen messen* (1. Auflage). Stuttgart: Schäffer-Poeschel Verlag.
- Schüller, K., Busch, P. & Hindinger, C. (2019). Future Skills: Ein Framework für Data Literacy. Kompetenzrahmen und Forschungsbericht. Hochschulforum Digitalisierung. Zugriff am 7.9.2020. Verfügbar unter: https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_Nr_47_DALI_Kompetenzrahmen_WEB.pdf
- Schulze-Krüdener, J. (2017). *Wissen, was in der Sozialen Arbeit wirkt! Zur Reichweite empirischer Zugänge* (1. Auflage.). Bremen: Apollon University Press.
- Wachsmuth, E., Brinkmann, H. & Hense, J. U. (2019). Wirkungsorientierung in der Non-Profit-Programmarbeit – Wirkungsmodelle und Wirkungsmodelltypen als Instrumente der Steuerung und Evaluation. *Zeitschrift für Evaluation*, 18(2), 291–306. <https://doi.org/10.31244/zfe.2019.02.07>

Carolin Freier

„Die Zahlen stehen jedoch wieder drüber“

Nichtintendierte Folgen einer Indikatoren gestützten Wirkungsorientierung in der Beratung

In der Berufspraxis stehen Beschäftigte in der Sozialen Arbeit sowie der Sozialwirtschaft häufig vor der Herausforderung, aus dem vorhandenen Wissen ein adäquates Vorgehen im konkreten Fall abzuleiten. Diese Wahl kann durch die Vielzahl der auf unterschiedlichen theoretischen Vorannahmen beruhenden Methoden (z. B. Galuske 2013; Kreft, 2017, Müller, 2017; Wendt, 2017; Schermer, Weber, Drinkmann & Jungnitsch, 2016) erschwert werden, auch wenn Methoden teils auf spezifische Handlungsfelder zugeschnitten sind (z. B. Kilb & Peter; 2016, für den schulischen Bereich). Doch die sozialarbeiterische Praxis ist komplex und es besteht ein Bedarf an Orientierung und Handlungssicherheit für Beschäftigte in der Sozialen Arbeit, z. B. wie sie mit ihren Adressat*innen in der aktuellen Situation umgehen sollen, welche Instrumente sich im konkreten Anlass bereits ‚bewährt haben‘ oder welches Handeln notwendig respektive verzichtbar ist. Wirkungsanalysen genauso wie „Sozialforschung von unten und in eigener Sache“ (Riemann, 2004, S. 190) bieten Antworten auf solche Fragen: Durch Methoden der empirischen Sozialforschung kann evidenzbasiertes Wissen für wirksames Handeln bereitgestellt und das professionelle Handeln in der Praxis systematisch überprüft werden (siehe z. B. für die evidenzbasierte Urteilsbildung Bastian, 2019). Ziel sollte dabei sein, dass die Wirkungsanalyse Sozialarbeiter*innen bei der „operativen Arbeit, aber auch bei ihrer strategischen Planung hilft“ und nicht allein für das Management erstellt wird (Bono, 2019, S. 26). Diesem Verständnis folgend sollten Wirkungsanalysen in der Sozialen Arbeit genutzt werden, um die Wirksamkeit professionellen Handelns zu analysieren. Dabei bemisst sich Wirkung im Selbstverständnis der Freien Wohlfahrtsverpflege an einer möglichst effektiven „Verbesserung von hoch individuellen Lebenslagen sowie die Sicherung bzw. Ermöglichung von gesellschaftlicher Teilhabe – ein Rechtsanspruch, der in den Sozialgesetzbüchern geregelt ist“ (Fridrich & Timm, 2019, S. 28). Zum anderen wird eine Wirkungskontrolle mitunter auch gesetzlich von Leistungserbringenden eingefordert, etwa im Bundesteilhabegesetz (§§121, 125, 128 SGB IX). Beispielsweise ist Social Return on Investment (SROI) ein Ansatz, der wachsendes Interesse als Wirkungsanalyse erfährt (Schober & Then, 2015, S. 3). Die Fachdebatte zu Wirkungsanalysen liefert dabei ein polarisierendes Bild: Indikatoren, die Professionskriterien-entsprechendes Handeln abbilden, beinhalten die Chance, Wirkungen zu erfassen. In Folge kann professionelles Handeln daran ausgerichtet werden, was unter welchen Umständen wie wirkt. Dabei ermöglichen Wirkungsanalysen nicht nur die Sichtweise der Profession oder des Managements abzubilden, sondern ebenso die Einschätzung der Adressat*innen sowie unterschiedlicher beteiligter Beschäftigter zu Wirkungen direkt einzubeziehen. Wirkungsanalysen können dann dazu beitragen eine breitere Perspektive auf Wirkungen zu etablieren und auch den Wert von Nonprofit-Aktivitäten aufzuzeigen. Chancen der langfristigen Qualitätsverbesserung Sozialer Arbeit sind demnach zu erwarten, wenn für die jeweiligen Arbeitsbereiche individuelle und praxistaugliche

Verfahren der Wirkungserfassung entwickelt werden (BAGFW, 2015; Bono, 2019; Friedrich& Timm, 2019; Ottmann & König, 2019). Dementsprechend wird argumentiert, dass basierend auf Wirkungsanalysen die Ressourcenallokation effektiver und effizienter erfolgt.

Hingegen werden Wirkungsanalysen kritisiert, wenn die Bewertung komplexer Zusammenhänge ungerechtfertigt vereinfacht und dementsprechend unzulänglich gemessene Wirkungen objektiviert werden (Frey, 2007). Dabei bleiben normative Vorprägungen der Wirkungsdefinition intransparent, etwa in Form von Gewichtungen (Bibisidis, 2018). So sind die Zieldefinitionen, Methoden der Wirkungserfassung und deren Ergebnisinterpretation bedeutsam für die Ermittlung und Beurteilung von Wirkung.

Neben methodischen Fragen agiert eine an Wirkungen orientierte Soziale Arbeit im Rahmen eines sozialstaatlichen Settings (siehe auch Otto, Wohlfahrt & Ziegler, 2019). Vor dem Hintergrund finanzieller und wirksamkeitsorientierter Fragen führen sozialstaatliche Regulierungen bereits seit den 1980er Jahren zu einem neuen Verhältnis von Staat und Wohlfahrtsverbänden, in dem Wohlfahrtsverbände in Wettbewerb mit privaten Unternehmen treten und an Effizienzkriterien bemessen werden (Schroeder, 2019, S. 26). Die Wirkungsorientierung in den Steuerungsansprüchen des Staates ist auch anhand der Einführung konkreterer staatlicher Vorgaben und Kontrollen rekonstruierbar (Schroeder, 2019). Weiter ist die Wirkungsorientierung verbunden mit Implikationen für die Steuerung in wohlfahrtsstaatlichen und privaten Organisationen, die Soziale Arbeit erbringen. Die Erfassung von Wirkungen trägt dazu bei, dass sich eine Wirkungsorientierung reproduziert. Denn Wirkungsanalysen generieren Daten und Indikatoren basierte Formen der Bewertung sowie Kontrolle Sozialer Arbeit, sozialwirtschaftlicher Einrichtungen und Verwaltungen. Erfasst in Monitoring- und Steuerungssystemen, dienen solche Formen der Bewertung als Grundlage managerialer Entscheidungen. Damit haben rekonstruierte Wirkungen nicht nur Einfluss auf das Bild erbrachter sozialarbeiterischer Leistungen, sondern prägen in Form organisationaler Regulierungen auch Rahmenbedingungen des professionellen Handelns in direkter Interaktion mit den Adressat*innen. Fragen nach Effizienz und ökonomischen Wirkungen können dabei dem Ziel der Verbesserung individueller Lebenssituationen und sozialer Teilhabe von Menschen entgegenstehen (Burmester& Wohlfahrt, 2018).

Hier setzt der vorliegende Beitrag¹ an, der zur Fachdebatte in der Sozialarbeitswissenschaft beitragen möchte, indem er empirische Ergebnisse zu nichtintendierten Folgen wirkungsorientierter Steuerung für die Handlungspraxis von Berater*innen in der Arbeitsvermittlung vorstellt. Mit dem Fokus auf Arbeitsvermittlung behandelt der Beitrag zum einen das auch für die Soziale Arbeit relevante Handlungsfeld der Qualifizierung und Beschäftigung (Freier, 2020). Zum anderen ist die Bundesagentur für Arbeit ein Vorreiter in der an Wirkungen in Form von Zielindikatoren orientierten Steuerung, die bis in das Case Management in der Arbeitsvermittlung hineinreicht (Freier & Senghaas, 2021; Sowa & Staples, 2014; Freier, 2016, S. 155ff.). Die Analyse dieses Beratungssettings ermöglicht Rückschlüsse auf potenzielle Veränderungen in der Steuerung von Trägern Sozialer Arbeit mit einer durchgreifenden Wirkungsorientierung im Sozialinvestitionsstaat.

Dafür wird zunächst die Bedeutung von Wirkungen im sozialstaatlichen Handeln als Sozialinvestitionsstaat skizziert (Abschnitt 1). Dann werden die Arbeitsvermittlung und die Rolle

¹ Ich danke Markus Gottwald und Katja Hartosch für wertvolle Hinweise.

der Wirkungsorientierung in der Bundesagentur für Arbeit vorgestellt (Abschnitt 2). Darauf aufbauend wird die Praxis der Arbeitsvermittlung mit Blick auf die wirkungsorientierte Steuerung analysiert. Genutzt werden Evaluationsergebnisse eines Innovationsprozesses durch das Institut für Arbeitsmarkt- und Berufsforschung in Kooperation mit dem Fachbereich Wirkungsanalyse der Bundesagentur für Arbeit (Freier & Senghaas, 2021; Freier, Kupka, Senghaas & Wuppinger, 2017). Im Fazit (Abschnitt 4) wird die Wirkungsorientierung mit Blick auf professionelles Handeln und dessen Abbildbarkeit diskutiert.

1. Wirkungsorientierung im Sozialinvestitionsstaat

Die Sozialstaaten der westlichen Industrienationen entwickelten sich seit dem Ende des 20. Jahrhunderts hin zum Modell des Sozialinvestitionsstaats (Bothfeld & Rosenthal, 2018; Dwyer, 2004) – auch diskutiert unter den Begriffen „welfare to work“ (Wright, 2012) oder „aktivierender Sozialstaat“ (z. B. Lessenich, 2008; Serrano Pascual, 2007). Mit der Transformation zu einem Sozialinvestitionsstaat verschiebt sich der Schwerpunkt der Sozialleistungen weg von der Statussicherung hin zu sozialen Investitionen (Giddens, 1998). Während der traditionelle Sozialstaat versucht, die bestehende Soziallage seiner Bürger*innen durch Sozialtransfers zu stabilisieren und diese so vor dem Markt zu schützen, ist das Leitbild des Sozialinvestitionsstaats seine Bürger*innen zu stärkeren Akteur*innen im Markt selbst zu machen. In diesem Sinne sollen sozialpolitische Interventionen die Bürger*innen aktivieren, selbst tätig zu werden. Ein Beispiel dafür sind die Reformen in der Sicherung bei Arbeitslosigkeit: hier greifen zum einen Entgeltersatzleistung aus der Arbeitslosenversicherung und zum anderen die bedarfsgeprüfte und steuerfinanzierte Grundsicherung für Arbeitsuchende. Insbesondere die Ausgestaltung der Grundsicherung für Arbeitsuchende als Grundsicherung und die Priorisierung der Leistungen aktiver Arbeitsmarktpolitik (vor dem passiven Leistungsbezug) binden Sozialleistungen an nachweisbare Eigenbemühungen von Arbeitsuchenden rück (Freier, 2016, S. 47ff.). Sozialleistungen werden also im Sozialinvestitionsstaat konditionalisiert (Serrano Pascual, 2007) und führen zu vielfach diskutierten sozialen Verwerfungen (z. B. Offe, 1998; Leitner & Lessenich, 2003; Bothfeld & Rosenthal, 2018).

Sozialleistungen werden dabei nicht allein für Adressat*innen konditionalisiert. Betrachtet als soziales Investment wird die Verteilung von Ressourcen auch für sozialwirtschaftliche Träger und Erbringer von Sozialleistungen konditionalisiert.

„Die Aufgabe der lokalen Daseinsvorsorge, Güter und Dienstleistungen für Alle bedarfsgerecht zur Verfügung zu stellen, wird in ein Mikromanagement von Sozialunternehmen zum Zweck einer Optimierung des gesamtgesellschaftlichen Nutzens neu definiert, ebenso der gesamtgesellschaftliche Nutzen als monetäre Kosten-Nutzen-Bilanz von Wohlfahrtsausgaben“ (Otto, Wohlfahrt & Ziegler, 2019, S. 92).

Infolge kommt Wirkungsanalysen im Sozialinvestitionsstaat auch die Funktion zu, Daten zu generieren, die Entscheidungen für den Ressourceneinsatz ermöglichen. Dementsprechend werden in der Fachdebatte Sozialer Arbeit zu Recht Risiken der mit Wirkungsanalysen einhergehenden Legitimierungserwartungen geäußert, die die Ökonomisierung der Sozialen Arbeit im Sozialinvestitionsstaat verschärfen (z. B. Otto & Wohlfahrth, 2019; Ziegler & Wohlfahrth,

2019). Folglich liefern Wirkungsanalysen die benötigte Datengrundlage für selektive Investitionen – pay for result and pay for success (Burmester & Wohlfahrt, 2018), die mit Blick auf Effizienz- und Effektivitätskriterien Investitionen gerade dort ermöglichen, wo Wirkungen abbildbar sind respektive abgebildet werden.

Um Arbeitsprozesse und deren Wirkungen abbilden zu können, ist die Generierung von Prozessdaten ein probates Mittel, die in Monitoring- und Steuerungssystemen erfasst werden. Welche Wirkungen wie erhoben werden, prägt dabei das Bild von der Art und dem Umfang der erbrachten Leistungen. Basierend auf dieser Leistungs- und Wirkungsbemessung werden typischerweise Entscheidungen zur Steuerung der Leistungen getroffen, die sich wiederum bis auf die Ebene der Arbeit mit den Adressat*innen auswirkt. Bei Trägern der Eingliederungshilfe nach dem neuen Bundesteilhabegesetz führt dies bereits zur „Erstellung und Anwendung ökonomisch-schematisierter Hilfeplan- und Verlaufsmodelle und damit verbundener objektivierbarer, statistisch messbarer Ergebnisindikatoren etwa in Form von Kennzahlen“ (Kratz, 2017, S. 33). Zusammengefasst bilden, durch Wirkungsanalysen generierte Daten und Indikatoren basierte Formen der Bewertung und Kontrolle erbrachter sozialstaatlicher Leistungen, eine bedeutende Entscheidungs- und Steuerungsgrundlage für den Sozialstaat.

2. Wirkungsorientierung in der Bundesagentur für Arbeit

Um zu analysieren, wie sich die Wirkungsorientierung wiederum auf die Praxis auswirkt, ist die Arbeitsvermittlung in der Bundesagentur für Arbeit (BA) ein geeignetes Fallbeispiel. Die Behörde ging aus einem Restrukturierungsprozess der Bundesanstalt für Arbeit hervor. Öffentliche Aufmerksamkeit unter dem Stichwort „Vermittlungsskandal“ erregten gravierende Fehler in der Vermittlungsstatistik, die der Bundesrechnungshof im Jahr 2002 feststellte. Wie in fast allen europäischen Staaten wurde New Public Management im Zuge der Restrukturierung zu einer zentralen Säule der Beschäftigungspolitik in Deutschland (Brülle, Krätschmer-Hahn, Reis & Siebenhaar, 2016, S. 15). Im Sinne einer managerialen Steuerung wurden eine Ergebnis- und Wettbewerbsorientierung in der Organisationsform der BA verankert, mit dem Ziel, die öffentliche Verwaltung effektiver und kostengünstiger zu erbringen. Dafür wurde erstens ein umfangreiches Indikatoren basiertes Monitoringsystem etabliert. Es bemisst die Wirkungen der Arbeitsvermittlung beispielsweise an der Integrationsquote, der Dauer der Arbeitslosigkeit, den erfolgreich besetzten Stellen oder der kumulierten Beschäftigungsdauer und dokumentiert die Arbeit der Berater*innen mit Blick auf die Einhaltung von Prozessvorgaben in Kennzahlen. (Freier et al., 2017). Zweitens steuert die BA kaskadenförmig die Umsetzung politischer Zielsetzungen in den leistungserbringenden Agenturen für Arbeit: ausgehend von mehrjährigen Rahmenzielen, die das Bundesministerium für Arbeit und Soziales mit der BA vereinbart, schließt der BA-Vorstand jährliche Zielvereinbarungen mit den Geschäftsführungen der Regionaldirektionen, die ihrerseits mit Geschäftsführungen der Agenturen für Arbeit Zielvereinbarungen abschließen. Die Indikatoren gestützte Zielsteuerung ermöglicht dabei nicht nur die Abbildung der Zielerreichung einzelner Agenturen für Arbeit, sondern auch deren Vergleich mit anderen Agenturen für Arbeit mit ähnlichen regionalen Arbeitsmarktbedingungen, um auf Basis der Kennzahlen Wettbewerb zu inszenieren (Brülle et al., 2016, S. 19). Auch innerhalb der Agenturen für Arbeit werden die Zielindikatoren für teamweise Vergleiche herangezogen (Sowa & Staples, 2014).

Dabei unterstellt die Verwendung der Wirkungsindikatoren im Kontext des internen Wettbewerbs, dass diese tatsächlich die organisationale Realität so abbilden, dass „unterschiedliche qualitativ definierte Zustände quantitativ miteinander verglichen werden können“ (Brülle et al., 2016, S. 19).

3. Empirische Schlussfolgerungen zu Folgen von Wirkungsorientierung in der Arbeitsvermittlung
Welche nichtintendierten Folgen die Wirkungsorientierung als Element von New Public Management für professionelle Arbeit mit sich bringt, wird am Fallbeispiel der wirkungsorientierten Steuerung in der Handlungspraxis von Berater*innen in der Arbeitsvermittlung dargestellt. Herangezogen werden dafür Ergebnisse der Begleitforschung eines Innovationsprojektes in der Arbeitsvermittlung, das auf die Weiterentwicklung interner Arbeitsprozesse und Dienstleistungen durch die Partizipation der Beschäftigten zielte (Freier et al., 2017). Während typischerweise über die Zentrale der BA standardisierte Verfahren die Beratungssituation zwischen Berater*innen und Arbeitssuchenden respektive Arbeitgeber*innen prägen, hatte ein in drei Agenturen für Arbeit laufendes Innovationsprojekt das Ziel, in den Jahren 2015 und 2016 von verwaltungsinternen Richtlinien, Weisungen und Prozessen abweichen zu können. Unverändert blieben dabei im Sozialgesetzbuch III definierte Ziele (Vorrang der Vermittlung und aktiver Arbeitsförderung, Leistungen wie Maßnahmen zur Aktivierung und beruflichen Eingliederung), Grundstrukturen der Arbeitsorganisation, die Zuteilung von Haushaltsmitteln und die Messung von Zielindikatoren im Zielsystem. Die Agenturen für Arbeit betreuen als Teil der Arbeitslosenversicherung jene Arbeitssuchende, die weniger als ein Jahr arbeitslos sind. Seitens der BA war mit dem Projekt die Erwartung verbunden, die Arbeitsmarktintegrationsleistung durch Innovationen zu steigern. Daher wurden explizit keine zusätzlichen Ressourcen für den Innovationsprozess bereitgestellt, die einen höheren messbaren Output generieren könnten. Für die Analyse nichtintendierter Folgen der wirkungsorientierten Steuerung ist das Projekt insbesondere daher geeignet, da teilnehmende Agenturen für Arbeit im Rahmen des ersten Projektjahres die mit den Regionaldirektionen vereinbarte Zielerreichung (ausgedrückt in Zielindikatoren) weniger ernst nehmen sollten, um der Kreativität und Innovation der Berater*innen Raum zu geben.

Nichtintendierte Folgen der Wirkungsorientierung auf die professionelle Arbeit von Berater*innen lassen sich am Beispiel der Innovationen in der Vermittlungsstrategie illustrieren, die im Rahmen von Interviews mit Fach- und Führungskräften sowie Gruppendiskussionen erhoben wurde (für weitere Ergebnisse siehe Freier et al. 2017; Freier & Senghaas, 2021).

In den Agenturen für Arbeit werden Arbeitssuchenden Stellenangebote ausgegeben. Neben dem Austausch von Berater*innen mit ihren Adressat*innen laufen im Hintergrund digitale Matching-Prozesse zwischen Anforderungen offener Stellenangebote und den Kompetenzen der Arbeitssuchenden ab. Dabei sind Beratungsprozesse für die Arbeitsvermittlung in Agenturen für Arbeit nach Arbeitsmarktseiten getrennt:

Berater*innen und die von ihnen betreuten Arbeitssuchenden hinterlegen im IT-System der BA VerBIS berufliche Qualifikationen und Kompetenzen sowie Informationen zur Mobilität (etwa, ob ein Führerschein vorhanden ist). Diese Berater*innen haben typischerweise keinen Kontakt zu

Arbeitgeber*innen.² Auf der anderen Seite hinterlegen Berater*innen, die Arbeitgeber*innen betreuen, Stellenangebote und dafür benötigte Kompetenzen im IT-System. Arbeitsuchende erhalten dann ein Stellenangebot, den sogenannten Vermittlungsvorschlag, entweder von ihrer*em zuständigen Berater*in oder von Berater*innen, die Arbeitgeber*innen betreuen.

Im Zuge des Innovationsprojektes wurden die seit der Organisationreform getrennten Bereiche der arbeitgeber- und der arbeitnehmerorientierten Vermittlung stärker miteinander verschränkt. Während zuvor die Schnittstelle insbesondere im IT-basierten Matching-Verfahren lag, das Stellengesuche passenden Arbeitsuchenden zuordnete, wurden im Projekt die Kenntnisse beider Arbeitsmarktseiten über formelle oder informelle Austausch der Berater*innen intensiviert, teamweise zusammengearbeitet oder die Betreuung beider Arbeitsmarktseiten von spezifischen Berater*innen in einer Person übernommen. So verstärkten Berater*innen Vermittlungsarbeit im direkten Austausch von Arbeitsuchenden und Arbeitgeber*innen. Beispielsweise organisierten Berater*innen ein Kennenlernen mehrerer Arbeitsuchender mit potentiell passenden Arbeitgeber*innen und begleiteten die Arbeitssuchenden dahin. Ferner wurden Bewerbungsgespräche im Rahmen von Informationsveranstaltungen in der Agentur für Arbeit veranstaltet. Besondere Aufmerksamkeit wurde der Vermittlung von eher arbeitsmarktfernen Arbeitsuchenden gewidmet. Hier sollten passende Stellenangebote eruiert werden, die Gemeinsamkeiten von Kompetenzen der Arbeitsuchenden und Stellenanforderungen beinhalten, allerdings vom IT-gestützten Matching-Verfahren weniger erfasst werden. Weiter wurde mit Arbeitgeber*innen verstärkt abgestimmt, ob bei spezifischen Stellenanforderungen Kompromisse möglich sind. Eine zentrale Strategie war hier, sich auf wenige qualitativ hochwertige Vermittlungsvorschläge zu konzentrieren, die eine gute Passung zwischen Arbeitsuchenden und Arbeitgeber*innen herstellen, statt eine quantitativ umfangreichere Ausgabe von Vermittlungsvorschlägen vorzunehmen, die weniger mit den Beteiligten abgestimmt wurden. Folgendes Zitat einer Fachkraft bringt diese Strategie auf den Punkt: „also der Kunde³ steht einmal im Vordergrund, auch unser Metaziel, wir wollen Menschen in Arbeit bringen und Arbeitgeber und Bewerber zusammenführen und nachhaltig integrieren“⁴ (85, 173). Die zuvor geltende Annahme, je mehr Stellenangebote ausgegeben werden, desto mehr Arbeitsuchende finden eine Arbeit, wurde damit auf den Prüfstand gestellt.

Doch mit dem Einbruch des Indikators für die Arbeitsmarktintegration wurde die Projektstrategie, sich auf eine hohe Qualität der Vermittlungsvorschläge zu konzentrieren, aufgegeben und stattdessen wieder deren Anzahl erhöht (Freier & Kupka, 2016).

Dass dabei nicht nur die Projektrahmung verletzt wurde, im ersten Jahr der Zielerreichung im Rahmen der wirkungsorientierten Zielsteuerung eine geringere Bedeutsamkeit zu verleihen, sondern auch auf der Ebene der Beratung Zielindikatoren eine hohe Bedeutung im Arbeitsalltag einnehmen, führen folgende Stimmen aus einer Gruppendiskussion vor Augen:

² Die Trennung in arbeitgeber- und arbeitnehmerorientierte Vermittlung findet sich auch in Jobcentern. Ausnahmen existieren hier teilweise im Fallmanagement mit Personen mit besonderem Vermittlungsbedarf.

³ In Agenturen für Arbeit werden die Adressat*innen als „Kund*innen“ bezeichnet (zur Diskussion, siehe Freier, 2016, S. 52ff.).

⁴ Die Zitate werden für eine bessere Lesbarkeit sprachlich geglättet. Zitate werden mithilfe der Interviewnummer und dem Absatz in MAXQDA nachgewiesen.

*Berater*in 1: Die ersten acht Monate, hat man uns das Gefühl gegeben, das läuft alles relativ locker, Zahlen spielen keine Rolle. Jetzt dreht sich das total – auf einmal sind die Zahlen in jeder Besprechung, in jedem Teamboard, in jeder Teambesprechung und auch (etwas lachend) zwischendrin. Die Zahlen stehen ganz groß [...]*

*Berater*in 2: Die Zahlen stehen jetzt doch wieder drüber, ne?*

*Berater*in 1: [...] im Vordergrund. (29: 175ff.)*

Die Strategie, sich auf qualitativ hochwertige Vermittlungsvorschläge zu konzentrieren und insbesondere das Wissen zu arbeitgeber- und arbeitnehmerorientierter Vermittlung zu erweitern, um ein individuell passendes und längerfristiges Vermittlungsergebnis zu erreichen, wurde damit nicht im Projektverlauf beibehalten. Vermittelt über Teambesprechungen und Gespräche mit Vorgesetzten erfahren die Berater*innen, dass die unter ihrer Beteiligung entwickelten Maßnahmen zur Verbesserung der Qualität in Vermittlung und Beratung zunächst ausgesetzt werden müssen, um beeinflussbare Zielindikatoren in Richtung einer Erfolgsabbildung zu steigern. Obwohl auch für die nachhaltige Integration ein Zielindikator in die Beschäftigungsstatistik eingeht, der angibt, ob Arbeitsuchende mindestens sechs Monate beschäftigt sind, werden im Projektkontext Zielkonflikte für die Berater*innen deutlich.

Zwar würden bei der Fokussierung auf eine nachhaltige Entwicklung beispielsweise passgenauere, in der Summe aber weniger Vermittlungsvorschläge ausgegeben. Doch dienen im wirkungsorientierten Zielsteuerungssystem die Anzahl der Vermittlungsvorschläge als quantifizierbarer Nachweis von Vermittlungsaktivitäten. Hinzu kommt, dass der Indikator einer erfolgreich besetzten Stelle nur dann verbucht wird, wenn vorab ein Vermittlungsvorschlag ausgestellt wurde. Dieses Verfahren ist durchaus einsichtig, um direkte Wirkungen der Arbeitsvermittlung zu erfassen. Doch in der Handlungspraxis führt die Vermessung des Handelns der Berater*innen dazu, dass die Projektstrategie verändert wurde (Freier & Senghaas, 2021).

Die dargestellte Kehrtwende in der Vermittlungsstrategie ist insbesondere darauf zurückzuführen, dass sich innerhalb des Projektes der Umgang mit der Zielerreichung veränderte: galt zu Projektbeginn die Erreichung der Zielvorgaben im ersten Jahr als irrelevant, wurde diese Rahmung im Projektverlauf zurückgenommen. Top-Down werden die „Zahlen“ über die verschiedenen Ebenen an die Berater*innen weitergegeben, und der Druck verstärkt, dass die Agentur für Arbeit zur Zielerreichung beiträgt.

Über dieses Projektspezifikum hinausgehend, zeigt jedoch die Diskussion der Berater*innen, dass diese in Teamsitzungen und im Gespräch mit Vorgesetzten häufig mit jenen Zahlen konfrontiert sind, die die Wirkung ihrer Arbeit messen sollen. Auch wenn weder die Berater*innen noch deren Führungskräfte an der Zielerreichung bemessene Leistungsgratifikationen, leistungsbasierte Entgelte oder ähnliches erhalten, besitzen „die Zahlen“ eine hohe Relevanz in der Organisation (für weitere Ergebnisse zu Einflüssen des Zielsystems auf das Handeln der Berater*innen werden Ergebnisse eines entsprechenden Forschungsprojektes mit Spannung erwartet, siehe zum Forschungsdesign Sowa, Gottwald, Grimminger, Ixmeier & Promberger, 2016).

Doch ist die Abbildbarkeit von Wirkungen begrenzt, was eine Bereichsleitung für die arbeitgeberorientierte Vermittlung folgendermaßen pointiert: „wir haben Zielwerte als Agentur, das sind aber erst mal Werte, Sequenzen von einem großen Prozess, die nur über diesen Indikator gemessen werden“ (19, 59). Die Dienstleistung sei jedoch weitreichender und solle als eigentliches Ziel in den Blick genommen werden. Daher benötigt die wirkungsorientierte Zielsteuerung eine spezifische Kultur der Interpretation von Indikatoren, für die die Bereichsleitung Unterschiede hinsichtlich der Führung der lokalen Agentur für Arbeit auf der einen Seite gegenüber der Regionaldirektion und Zentrale der Bundesagentur für Arbeit auf der anderen Seite festmacht:

„also wir [lokale Agentur für Arbeit – C.F.] differenzieren das glaub ich sehr, sehr gut, was sind Ziele und was ist ein Indikator und welche Sequenz dieses Ziels wird überhaupt nur gemessen. Aber in dem Erleben wie wir geführt werden, werden wir oft reduziert auf die Indikatoren.“ (19, 59)

Trotz eines ausdifferenzierten Zielsteuerungssystem, dessen Indikatoren – nicht allein im Controlling, sondern auch auf der Leitungsebene – im Wissen methodisch-begrenzter Aussagekraft interpretiert werden, können die wirkungsorientiert produzierten „Zahlen“ eine eigene Wirkmächtigkeit entfalten, die sich von deren methodischer Erfassung loslösen. Auf den unterschiedlichen Hierarchieebenen müssen sich dann Regionaldirektionen, Agenturen für Arbeit und Teams für die im Zielsteuerungssystem abgebildeten Wirkungen legitimieren. So wird deutlich, dass die Orientierung an Zielindikatoren im Rahmen einer wirkungsorientierten Steuerung die Gefahr beinhaltet, Eigensinn und individuelle Bedarfe der Arbeitsuchenden in nicht abbildbaren Integrationsschritten als blinde Flecken nicht abzubilden (Bender & Brandl, 2017) und damit auch professionelle Arbeit von Berater*innen nicht sichtbar zu machen (Freier & Senghaas, 2021).

4. Fazit

Am Beispiel der Arbeitsvermittlung analysierte der Beitrag, inwiefern über Daten- und Indikatoren basierte Formen der Bewertung von Wirkung, professionelles Handeln in der Beratung beeinflusst werden kann. Dass im Rahmen eines Innovationsprojektes anvisierte Ziel, durch abgestimmtere Passung qualitativ hochwertige Stellenangebote zu platzieren, wurde aufgrund dessen mangelnder Erfolgsdarstellung in den Zielindikatoren zugunsten eines höheren quantitativen Volumens an ausgegebenen Stellenangeboten aufgegeben (Freier & Kupka, 2016).

Das vorliegende empirische Material zur wirkungsorientierten Zielsteuerung zeigt, dass neben der Diskussion um die methodisch valide und praxistaugliche Wirkungserfassung (Bono, 2019; Friedrich & Timm, 2019) insbesondere relevant ist, wie Ergebnisse der Wirkungserfassung in die Steuerung und Organisationskultur eingehen. Selbst mit methodisch aufwändiger Wirkungserfassung – in der Bundesagentur für Arbeit existieren mehrere Tausend Kennzahlen – ist die Komplexität von Prozessen nicht gänzlich abbildbar. Selbst wenn Leitungsebenen für die Interpretation notwendiges Kontextwissen (Bibisidis, 2018) darüber besitzen, welche Indikatoren welche Prozessequenzen abbilden und welche nicht, zeigt die zitierte Bereichsleitung der Agentur für Arbeit, dass in der Kommunikation mit weiteren Hierarchieebenen die Ergebnisinterpretation unzulänglich reduziert werden kann. Dabei lösen im Rahmen von

Indikatoren gemessene Wirkungen nicht allein Legitimationsdruck bei Zielverfehlung aus, sondern beeinflussen auch direkt die Beratung, so dass Berater*innen spüren: „Die Zahlen stehen jedoch wieder drüber.“

Unbestritten liefert eine wirkungsorientierte Kennzahlensteuerung hier Orientierung für Berater*innen in der Arbeitsvermittlung. Doch wie die vorgestellten Daten illustrieren, kann die Zielerreichung eine Eigendynamik erfahren. Zahlen erscheinen hier einen gewissen Selbstzweck einzunehmen im Kontext bestehender Macht- und Hierarchiestrukturen. Im beschriebenen Fall führen sie sogar zu einem Spannungsverhältnis zwischen professionellem Selbstverständnis von einer passenden an den Interessen und Bedürfnissen der Arbeitsuchenden ansetzenden gezielten Vermittlung hin zu einer weniger abgestimmten und umfangreicheren Ausgabe von Vermittlungsvorschlägen. Entsprechend der Abbildbarkeit der Wirkungen wird die Beratung hier wieder stärker formalisiert und standardisiert.

Die vorliegenden Analysen lassen den Schluss zu, dass auch in Organisationen Sozialer Arbeit Risiken bestehen, dass Zielindikatoren in der Steuerung und Organisationskultur nicht-intendierte Folgen mit sich bringen, wie eine Formalisierung mit Blick auf abbildbare Zielerreichung (siehe auch Bender & Brandl, 2017) und damit einhergehende Reduktion der Qualität Sozialer Arbeit (Seithe, 2019, S. 30). Diese Risiken verschärfen sich, wenn sich die Verteilung von finanziellen Ressourcen im Sozialinvestitionsstaat auf abbildbare Wirkungen konzentrieren.

Prof. Dr. Carolin Freier ist Professorin für Theorien und Handlungslehre in der Sozialen Arbeit an der Technischen Hochschule Nürnberg. Ihre Arbeitsschwerpunkte umfassen Veränderungsprozesse in Arbeit und Organisation, die Gestaltung des digitalen Wandels in Sozialer Arbeit sowie soziale Teilhabe und die Transformation des Sozialstaates.

Literatur

- Abadie, A., Diamond, A. & Hainmueller, J. (2010): Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490), 493–505.
- BAGFW (2015): Standortbestimmung der BAGFW zur Wirkungsorientierung in der Arbeit der Freien Wohlfahrtspflege. Berlin. Zugriff am 01.12.2019. Verfügbar unter <http://www.bagfw.de/uploads/media/bagfw-StandortbestimmungWirkungsorientierung.pdf>
- Bastian, P. (2019): *Sozialpädagogische Entscheidungen. Professionelle Urteilsbildung in der Sozialen Arbeit*. Stuttgart: UTB.
- Bender, G. & Brandl, S. (2017): Beschäftigungsorientierte Beratung im Spannungsfeld von Bürokratie und Professionalität. *Zeitschrift für Sozialreform* 63(1), 75–101.
- Bibisidis, T. (2018): Die aktuelle Debatte um die Wirkung (in) der Freien Wohlfahrtspflege. In: *Archiv für Wissenschaft und Praxis der sozialen Arbeit: Wirkungsorientierung in der Sozialen Arbeit*, 49(3), 38–45.
- Bono, L. M. & Verweijen-Slamnescu, R. (2019): „Wirkungsanalyse ist nicht Ziel, sondern Werkzeug“. *SOZIALwirtschaft*, 1, 25–27.
- Burmester, M. & Wohlfahrt, N. (2018): *Wozu die Wirkung Sozialer Arbeit messen?* Berlin: Lambertus.

- Bothfeld, S. & Rosenthal, P. (2018): The end of social security as we know it—The erosion of status protection in German labour market policy. *Journal of Social Policy*, 47(2), 275–294.
- Brülle, H.; Krätschmer-Hahn, R., Reis, C. & Siebenhaar, B. (2016): Zielsteuerung im SGB II. Kritik und Alternativen. Friedrich-Ebert-Stiftung. *Wisso Diskurs*, 09/2016, Bonn.
- Caliendo, M. & Hujer, R. (2005): *The Microeconomic Estimation of Treatment Effects – An Overview*. Bonn.
- Freier, C. (2016): Soziale Aktivierung von Arbeitslosen? Praktiken und Deutungen eines neuen Arbeitsmarktinstruments. (*Gesellschaft der Unterschiede*, 38), Bielefeld: transcript.
- Freier, C. (2020): Soziale Arbeit in der Beschäftigungsförderung. In: A. Rauch & S. Tophoven (Hrsg.), *Integration in den Arbeitsmarkt. Teilhabe von Menschen mit Förder- und Unterstützungsbedarf*, (*Grundwissen Soziale Arbeit* (S. 232–252)). Stuttgart: Kohlhammer.
- Freier, C. & Kupka, P. (2016): *Der „professionelle Fremde“ artikuliert das Unsagbare. Der Einfluss von Evaluation in der deutschen Arbeitsverwaltung*. Poster, präsentiert auf der 19. Jahrestagung der DeGEval, <http://dx.doi.org/10.13140/RG.2.2.26050.38084>
- Freier, C., Kupka, P., Senghaas, M. & Wuppinger, J. (2017): *Innovation und lokale Gestaltungsspielräume in der Arbeitsvermittlung. Begleitforschung zum Modellprojekt "Mach es einfach"*. (IAB-Forschungsbericht, 04/2017), Nürnberg.
- Freier, C. & Senghaas, M. (2021): Arbeitsvermittlung zwischen Nachhaltigkeit und Abbildbarkeit. Wie Vermittlungsfachkräfte Entscheidungsspielräume nutzen. *Soziale Welt* 72(2), 113–138. <https://doi.org/10.5771/0038-6073-2021-2-113>
- Freier, C. (i.E.): Metaphernanalyse als Methode in der Organisationsforschung. Ein Praxisbeispiel. In: R. Schmitt, J. Schröder, L. Pfaller & A.-K. Hoklas (Hrsg.), *Die Praxis der systematischen Metaphernanalyse*. Wiesbaden: VS, https://doi.org/10.1007/978-3-658-36121-1_12 (im Erscheinen).
- Frey, F. (2007): *Chancen und Grenzen von Wirkungsorientierung in den Hilfen zur Erziehung*. Wiesbaden: VS.
- Friedrich, P. & Timm, G. (2019): Wirkungserfassung in der Sozialen Arbeit – eine normative Diskussion, *SOZIALwirtschaft*, 1, 28–29.
- Galuske, M. (2013): *Methoden der Sozialen Arbeit: Eine Einführung*. Weinheim/München: Beltz Juventa.
- Giddens, A. (1998): *The Third Way. The Renewal of Social Democracy*. Cambridge: Polity Press.
- Kilb, R. & Peter, J. (2016): *Methoden der Sozialen Arbeit in der Schule*. München: Ernst Reinhardt.
- Kratz, D. (2017): Wie Wirkungsorientierung die Adressat_innen vergisst, *Sozial Extra* 17 (4), 32–34.
- Kreft, D. C. & Müller, W. (Hrsg.). (2017): *Methodenlehre in der Sozialen Arbeit: Konzepte, Methoden*. Erscheinungsort: Verlag.
- Lessenich, S. (2008): *Die Neuerfindung des Sozialen. Der Sozialstaat im flexiblen Kapitalismus*. Bielefeld: transcript.
- Leitner, S. & Lessenich, S. (2003): Assessing welfare state change: the German social insurance state between reciprocity and solidarity. *Journal of Public Policy* 23(3), 325–347.
- Offe, C. (1998): Der deutsche Wohlfahrtsstaat: Prinzipien, Leistungen, Zukunftsaussichten. *Berliner Journal für Soziologie* 8(3), 359–380.

- Otto, H.-U. & Wohlfahrt, N. (2019): Social Return on Investment: Anmerkung zum Antiprofessionalismus der aktuellen Wirkungsdebatte und ihres sachfremden Hintergrunds. *neue praxis*, 19 (2), 98-108.
- Otto, H.-U., Wohlfahrt, N. & Ziegler, H. (2019): Thesen zum ökonomischen Hintergrund einer verkehrten Konstruktion und ihren sozialpolitischen Folgen, *neue praxis*, 19 (2), 92-97.
- Ottmann, S. & König, J. (2019): Am Anfang steht das Wirkmodell, *Blätter der Wohlfahrtspflege*, 19 (2), 67-70.
- Riemann, G. (2004): Die Befremdung der eigenen Praxis, In A. Hanses (Hrsg.), *Biographie und Soziale Arbeit* (S. 241-260). Hohengehren: Schneider.
- Schermer, F., Weber, A., Drinkmann, A. & Jungnitsch, G. (2016): *Methoden der verhaltensorientierten Sozialarbeit: Basisstrategien*. Aschaffenburg: Alibri.
- Schober, C. & Then, V. (Hrsg.) (2015): *Praxishandbuch Social Return on Investment. Wirkung sozialer Investitionen messen*. Stuttgart: Schäffer-Poeschl.
- Schroeder, W. (2019): Die freie Wohlfahrtspflege und die wachsenden Steuerungsansprüche des Staates. *Sozialer Fortschritt*, 68 (1), 7-29.
- Seithe, M. (2019): Problematisches Professionsverständnis in der Sozialen Arbeit. *FORUMsozial*, 30-39.
- Serrano Pascual, A. (2007): Reshaping welfare states and activation regimes in Europe. In A. Serrano Pascual, L. Magnusson (Ed.), *Reshaping Welfare States and Activation Regimes in Europe* (pp.11-34). Brussels: Peter Lang
- Sowa, F. & Staples, R. (2014): Accounting in der Arbeitsverwaltung: Vermittlungsfachkräfte zwischen Steuerungsimperativen und autonomem Vermittlungshandeln. *Zeitschrift für Sozialreform*, 60(2), 149-173.
- Sowa, F., Gottwald, M., Grimminger, S., Ixmeier, S. & Promberger, M. (2016): *Vermittlerhandeln im weiterentwickelten Zielsystem der Bundesagentur für Arbeit. Zum Forschungsdesign einer organisationsethnografischen Studie* (IAB-Forschungsbericht 2/2016), Nürnberg.
- Wright, S. (2012): Welfare to work, agency and personal responsibility. *Journal of Social Policy* 41(2), 309-328.
- Ziegler, H. & Wohlfahrt, A. (2019): Zahlen und Geschichten – SROI als Umwertungstechnik im Sozialsektor, *neue praxis*, 19 (2), 109-124.

Sebastian Gogol

Soziale Arbeit als Investitionsobjekt? Private wirkungsorientierte Investitionen in Soziale Dienste

Zur Legitimation von Sozialausgaben, die oftmals auch als Investitionen in die Zukunft betrachtet werden, gewinnt die Frage nach der Wirksamkeit sozialer Dienstleistungsangebote zunehmend an Bedeutung. Als Investition betrachtete Sozialausgaben verlangen nach einer sozialen Rendite, die als Wirkungen auf bestimmte soziale Maßnahmen zurückzuführen und die auch monetär erfassbar sein sollen. So erfährt die Wirkungsdebatte in der Sozialen Arbeit eine neue, ökonomische Zielgröße und das Erfordernis von Wirksamkeit sozialer Maßnahmen ist mittlerweile expliziter Bestandteil sozialgesetzlicher Grundlagen (Bspw. BTHG). Eine wirkungsorientierte Steuerung sozialer Dienstleistungen stellt die Notwendigkeit zu erreichender Ziele und Ergebnisse in den Mittelpunkt der Betrachtung und definiert diese als Wirkung. So wird, anders als bei der Outputorientierung der Neuen Steuerung (Menge der sozialen Dienstleistungen), auf den Outcome (Ergebnisse und Zielerreichung) fokussiert. Der Prozess der Dienstleistungserstellung gerät somit in den Hintergrund und verliert als Kriterium für Qualität der sozialen Dienstleistung an Bedeutung (vgl. Burmester, Friedemann 2020: 208).

Social Impact Bonds (SIBs) greifen in ihrer Argumentationslogik die Definition von Sozialausgaben als Investment auf und stellen sich als ein neuartiges Finanzierungsinstrument für Soziale Dienste dar. SIBs sollen privaten Anleger*innen ermöglichen, Investitionen auf dem Sozialmarkt zu tätigen und damit innovative, wirkungsorientierte soziale Interventionen und Angebote vorzufinanzieren. Die Herstellung der Marktfähigkeit solcher Investments verlangt von Sozialen Diensten möglichst genau prognostizierbare und abbildbare Zielerreichungsgrade, die entsprechenden Methodeneinsatz und Wirkungsmessung erfordern. Monetär erfassbare Wirkungen stehen im Zusammenhang mit SIBs im Mittelpunkt der Betrachtung, da sich die finanzielle Rendite des Investments an zukünftigen Einsparungen oder erhofften zusätzlichen zukünftigen Steuereinnahmen oder einer Kumulation von beiden berechnet. SIBs sollen dazu beitragen (Re-)Finanzierung und Ausrichtung sozialer Arbeit grundlegend zu ändern und die Idee einer wirkungsorientierten und evidenzbasierten Sozialen Arbeit voranzutreiben (vgl. Burmester, Wohlfahrt 2018a: 93f). SIBs werden so zu einem sozialpolitischen Instrument, welches für die Soziale Arbeit weitreichende Folgen entfalten kann.

In Deutschland spielen SIBs für die Finanzierung Sozialer Dienste bislang eine eher untergeordnete Rolle. Wirkungsorientierung ist jedoch, wie bereits oben beschrieben, ein zentraler Bestandteil der Sozialgesetzgebung geworden und die Idee einer wirkungsorientierten und evidenzbasierten Sozialen Arbeit stößt auf zunehmendes Interesse verschiedener sozialpolitischer Akteur*innen. Insbesondere die Fragen, was als soziale Wirkung definiert werden soll und wie diese Wirkung zu messen sein kann sind umstritten.

Social Impact Bonds – Konstrukt und Herkunft

Social Impact Bonds sind eine sogenannte Multistakeholder-Partnerschaft. Es handelt sich dabei um eine Kooperation zwischen dem Staat oder gelegentlich auch einer philanthropischen Organisation als Auftraggeber, privaten Investor*innen und einem oder mehreren ausführenden Sozialunternehmen, die die soziale Maßnahme oder Intervention für eine bestimmte Zielgruppe

anbieten. Eine weitere Organisation oder Person fungiert als Intermediär*in, strukturiert den SIB, verhandelt Verträge zwischen den einzelnen Stakeholdern und beauftragt die unabhängige Evaluation der Maßnahme (vgl. Petrick, Weber 2013: 7). Als Wirkung werden zukünftige Einsparungen oder Steuereinnahmen für den Sozialstaat definiert, die bei Erreichen der zuvor festgelegten Ziele prognostiziert werden.

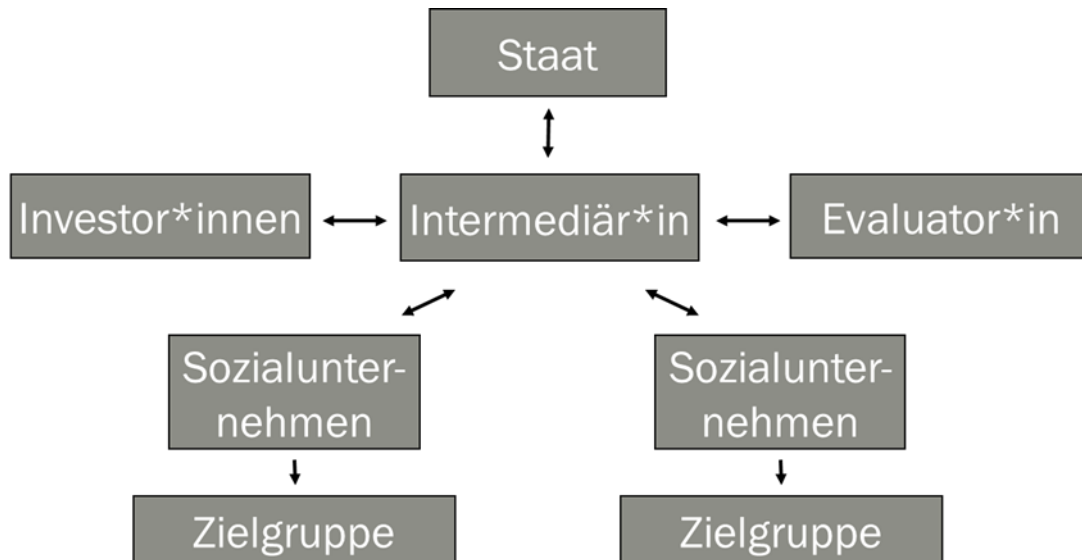


Abbildung 1: Struktur eines SIB (eigene Darstellung in Anlehnung an Fliegauß, 2014)

SIBs sollen also private Investitionen in eine soziale Wirkung ermöglichen, wobei die soziale Dienstleistung das Investitionsobjekt ist. Ein präventiver Ansatz der sozialen Intervention ist erforderlich, da es sich bei der angestrebten Kostenersparnis in der Regel um mögliche Folgekosten für den Sozialstaat bei Ausbleiben der Maßnahme handelt. Die Maßnahme selbst ist unter Umständen in der Durchführung kostenintensiver als vergleichbare regelfinanzierte Angebote. Die Kostenersparnis erfolgt also durch die angestrebte soziale Wirkung. Aus der prognostizierten Kostenersparnis errechnet sich dann eine finanzielle Rendite, die an die Investor*innen ausgezahlt wird. Nur bei Erreichen der vorab definierten Ziele (re-)finanziert der Staat die Maßnahme und schüttet ggf. die zuvor verhandelte Rendite aus. Es handelt sich hierbei um sogenannte „pay-by-results-Verträge“ oder auch „pay-for-success Finanzierung“. Das finanzielle Risiko liegt in diesem Konstrukt also bei den Investor*innen. Die Profitorientierung ist ein wichtiger Bestandteil von SIBs, wobei mindestens das eingesetzte Kapital als Rendite erwartet wird, aber durchaus auch marktübliche oder höhere Renditen verhandelt werden können (vgl. Petrick, Weber 2013: 5; Scheck, Weber 2012: 8).

Vor allem in Großbritannien sind Veränderungen in der Struktur von SIBs erkennbar. So übernimmt der Staat mittlerweile die Aufgaben des Intermediärs und führt selbst die Evaluation durch. SIBs werden dann für Sozialunternehmen ausgeschrieben, die sich selbst auf die Suche nach Investor*innen machen müssen und die entsprechenden Verträge direkt mit diesen verhandeln. Insbesondere die Aufgaben der Intermediär*innen und der unabhängigen Evaluation verursachen hohe Kosten, die SIBs unter Umständen nicht mehr rentabel erscheinen lassen, da die prognostizierten Einsparungen nicht zur Kostendeckung führen könnten. So werden auch in

anderen Fällen Intermediär*innen und Evaluation oftmals von Stiftungen finanziert. Zudem erfolgt mittlerweile häufig schon während der Laufzeit die Evaluation von Zwischenzielen und wird nicht erst nach Abschluss der Maßnahme durchgeführt. Dies ermöglicht die Auszahlung von Teilbeträgen und führt so zu einer Risikominimierung für Investor*innen. So reduziert sich zwar in der Regel die Gewinnaussicht, ein Totalverlust des eingesetzten Kapitals wird aber ausgeschlossen (vgl. Burmester, Wohlfahrt 2018b: 43ff).

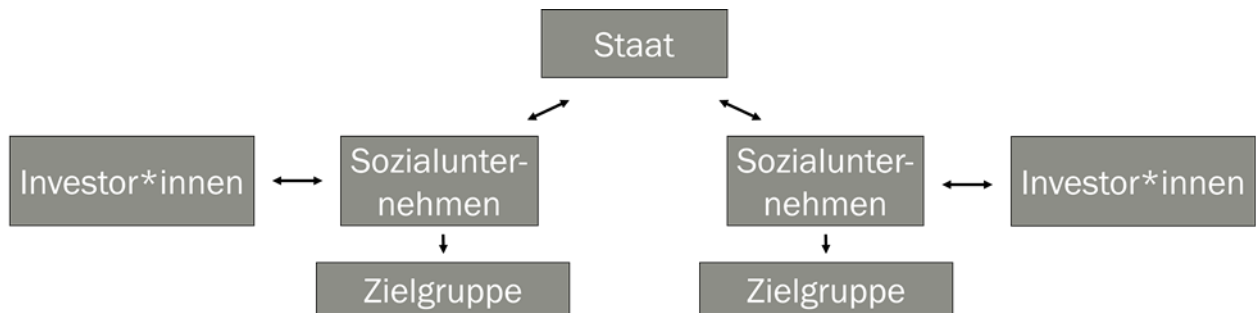


Abbildung 2: Strukturveränderung SIB

Der weltweite Einsatz von Social Impact Bonds geht auf eine Initiative aus dem Jahr 2013 des damaligen britischen Premierministers David Cameron zurück. Im Rahmen der britischen G8-Präsidentschaft initiierte er die Gründung der Social Impact Investment Taskforce (SIITF)¹. Sir Ronald Cohan, der bereits Vorsitzender der Social Impact Investment Taskforce in Großbritannien von 2000 bis 2010 war, übernahm den Vorsitz dieses Gremiums (vgl. NAB 2014: 4). Cohan sieht sich als ein Erfinder des Risikokapitals, des Social Impact Investments (Wirkungsorientiertes Investieren) in Großbritannien und letztlich auch der daraus entstandenen Synthese, der Social Impact Bonds. Im Jahr 2007 war er Mitbegründer von „Social Finance“^[2] in Großbritannien (vgl. Cohen 2014: 3).

Aus der Social Impact Investment Taskforce auf G8-Ebene wurden National Advisory Boards (NAB) initiiert, die in den jeweiligen G8-Staaten Bedingungen und Möglichkeiten für die Etablierung von Investments im Sozialbereich untersuchen und die Ergebnisse berichten sollten. Das NAB-Deutschland beschreibt in seinem Abschlussbericht die regelhafte staatliche Finanzierung, vor Allem im Bereich der Prävention und Innovation sozialer Maßnahmen und Interventionen, als nicht ausreichend, oder als erst gar nicht vorhanden. Da zudem gesellschaftliche Entwicklungen die Erschließung neuer Finanzierungsquellen erforderlich machen würden, sieht das NAB einen Bedarf an Maßnahmen zur Skalierung wirkungsorientierten Investierens (vgl. NAB 2014: 8).

¹ Dieses Gremium setzte sich aus staatlichen und zivilgesellschaftlichen Vertreter*innen der einzelnen Mitgliedsstaaten zusammen. Deutsche Vertreterinnen waren Susanne Dorasil, die Referatsleiterin des Bundesministeriums für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) und Dr. Brigitte Mohn von der Bertelsmann Stiftung als zivilgesellschaftliche Vertreterin.

² Social Finance ist eine Not-for-Profit-Organisation aus Großbritannien, die zur Förderung des wirkungsorientierten Investierens und der Sozialwirtschaft gegründet wurde. Seit 2010 tritt Social Finance regelmäßig als Initiator und Vermittler von Social Impact Bonds in Erscheinung. Mittlerweile gibt es in den USA, in Israel, Indien und den Niederlanden Schwesterorganisationen von Social Finance, welche als *Social Finance Global Network* miteinander kooperieren. Darüber hinaus finden sich weitere Netzwerk-Partnerorganisationen in Irland, Portugal, Südafrika, Canada, und den USA (vgl. Social Finance Global Network o.J.).

Die Lissabon Agenda aus dem Jahr 2000 definiert auf Ebene der EU Sozialausgaben als soziale Investitionen und gibt damit bereits zu diesem Zeitpunkt ein Denkmuster vor, welches Art und Weise von Sozialausgaben in Richtung (sozialer) Renditenerzielung lenken soll. Die Europäische Kommission forderte dann im Jahr 2013 die Mitgliedsstaaten auf, bei den Planungen für die EU-Fonds Sozialinvestitionen und "[...] die Prüfung innovativer Finanzierungsansätze und Finanzierungstechniken, die Analyse von Erfahrungen z. B. mit Social Investment Bonds [...]" (Europäische Kommission 2013: 27) mit einzubeziehen.

Der erste Kontinentaleuropäische Social Impact Bond wurde im Jahr 2013 in Deutschland aufgelegt. „eleven Augsburg/Jugendliche mit Perspektive (JuMP)“ sollte mindestens 20 Jugendliche und junge Erwachsene in sozialversicherungspflichtige Beschäftigung oder Ausbildung bringen. Ziel war der Erhalt der Beschäftigung für mindestens neun Monate. Die erhoffte Wirkung, also der gesellschaftliche Mehrwert, war hier demnach eine (Re-)Integration in den Arbeitsmarkt. Dieses Ziel deckt sich mit der Mehrzahl der derzeit international laufenden, abgeschlossenen und geplanten Social Impact Bonds (vgl. Burmester, Wohlfahrt 2018a: 87; GO Lab 2021). Im Jahr 2014 wurde der NAB-Deutschland Abschlussbericht vorgelegt, der eine sich weiter verschärfende, latente Unterfinanzierung des Sektors thematisiert und als Lösung private Investitionen vorschlägt: „Soziale, demographische und ökologische Entwicklungen bedingen die Suche nach neuen Finanzierungsquellen zur Lösung sozialer Herausforderungen“ (NAB 2014: 8).

Im September 2017 wurden zwei weitere SIBs in Deutschland aufgelegt „Prävention in den Hilfen zur Erziehung stärken – Im Landkreis Osnabrück“ und „Bildungschancen für Kinder verbessern in der Stadt Mannheim“. Beide sollen im folgenden Kapitel näher betrachtet werden.

Im März 2021 listet die *Impact Bond Global Database* von Social Finance 138 laufende und abgeschlossene, sowie 69 geplante SIBs, das *Impact Bond Dataset* des Government Outcomes Lab in Oxford international insgesamt 206 abgeschlossene, laufende und geplante Social Impact Bonds (vgl. GO Lab 2021; Social Finance UK 2021).

Social Impact Bonds in Deutschland – Mannheim und Osnabrück

Neben dem bereits abgeschlossenen SIB in Augsburg (hierzu ausführlicher vgl. Gogol 2019: 129ff) laufen derzeit zwei weitere SIBs in Deutschland. Beide wurden im Jahr 2017 aufgelegt, von der Bertelsmann Stiftung initiiert und die PHINEO gAG, einer Ausgründung aus der Bertelsmann Stiftung, betätigt sich in beiden SIBs als Intermediärin. Die Bertelsmann Stiftung beauftragt zudem die jeweilige Evaluation der Maßnahmen.

In Mannheim tritt die BASF SE als Investorin auf und erhält im Erfolgsfall die eingesetzten Gelder zurück. Dies soll den Vorteil haben, dass „[...] im Erfolgsfall zurückerhaltene Mittel erneut für Projekte zur Lösung gesellschaftlicher Herausforderungen eingesetzt werden können“ (Horning 2018: 4). Der SIB „Bildungschancen für Kinder verbessern in der Stadt Mannheim“ soll „herkunfts- und sozialbedingten Ursachen“ (ebd.) an der Grund- und Werkrealschule Pestalozzischule in Mannheim entgegenwirken. Unter der Überschrift Integrativer Campus Pestalozzischule soll neben Fortbildungen für Lehrer*innen, auch durch eine neue Unterrichtsgestaltung, Elternarbeit und Förderung der Kinder die Gymnasialempfehlungsquote gesteigert werden. Die Gymnasialempfehlungsquote ist in diesem SIB ein entscheidendes Wirkungskriterium. Wenn die „[...] herkunftsbedingte Benachteiligungen bei der Gymnasialempfehlungsquote signifikant

verringert werden konnten" (ebd.: 4) und „[...] das durchschnittliche kognitive Kompetenzniveau der Interventionsjahrgänge im Vergleich zu den Vorjahrgängen gesteigert werden konnte" (ebd.: 4) gilt der SIB als erfolgreich. Die Stiftung Fairchance, das Zentrum zur Therapie der Rechenschwäche, die Teach First Deutschland gGmbH und die KinderHelden gGmbH sollen im Zusammenwirken dazu beitragen diese Ziele zu erreichen.

Die Konzeption des Integrativen Campus ist allerdings nicht neu, sondern ähnelt sehr stark dem „Mannheimer Unterstützungssystem Schule – MAUS“, das bereits im Jahr 2008 von der Stadt Mannheim konzipiert und mehrere Jahre in Folge durchgeführt wurde (vgl. Stadt Mannheim o.J.: 2ff). Eine Innovation bezüglich der Maßnahme ist hier also nicht zu erkennen. Die Wirkungsmessung durch zusätzliche Lernstandtestungen könnte dazu führen, den Leistungsdruck bereits im Grundschulalter zu erhöhen und Konkurrenzdruck zu fördern. Es darf angezweifelt werden, ob dies als Mittel zur Förderung „[...] von sozialen und emotionalen Kompetenzen wie der Selbstwirksamkeit, der Kooperationsfähigkeit und der Verantwortungsbereitschaft" (Hornung 2018) sinnvoll ist. So ist die Innovation in diesem SIB letztlich auf die Finanzierungsform beschränkt und durch die jahrelange Erfahrung mit MAUS erscheint ein erfolgreicher Abschluss des SIBs als sehr wahrscheinlich. Erfolgreich in diesem Sinne ist dann allerdings nicht die erfolgreiche Präventionsarbeit, sondern die Generierung von Renditen, der Rückzahlung des eingesetzten Kapitals an die BASF SE. Die Wirkungsmessung durch die Gymnasialempfehlungsquote erscheint sehr subjektiv durch die beteiligten Lehrer*innen geprägt und kann zudem als eine nur kurzfristig erreichbare Wirkung betrachtet werden.

Im Landkreis Osnabrück startete ebenfalls im Jahr 2017 mit einer Laufzeit von drei Jahren der SIB „Prävention in den Hilfen zur Erziehung stärken – Im Landkreis Osnabrück“. Die Sparkasse Bersenbrück als Investorin erhält in diesem SIB eine zusätzliche Rendite nach erfolgreichem Abschluss des SIB, welche dann durch die Sparkasse Bersenbrück an eine Soziale oder eine Bildungseinrichtung gespendet werden soll. Das Sozialunternehmen Lega S gemeinnützige GmbH, ein anerkannter Träger der freien Jugendhilfe, führt die Maßnahme durch. Das Ziel ist die Vermeidung von Hilfen zur Erziehung für mindestens ein Jahr nach Abschluss der Maßnahme. Als Maßnahme kommt hier Triple P: Positive Parenting Program (dt.: Positives Erziehungsprogramm) zum Einsatz (vgl. Hornung 2017: 3f). Triple P ist ein zertifiziertes Trainingsprogramm für Eltern mit einem verhaltenstherapeutischen Ansatz. Es soll verhaltensbezogene und emotionale Probleme von Kindern positiv beeinflussen, indem es das Erziehungsverhalten von Eltern verändert. Es wird als evidenzbasiert beschrieben, da es laut Webseite von Triple P Deutschland auf „Über 800 Veröffentlichungen, klinische Untersuchungen, Wirksamkeits- und Disseminationsstudien rund um die Welt [...]“ (Triple P Deutschland GmbH o.J.) in den vergangenen 35 Jahren verweisen kann. Nachgewiesen ist hier für den Einsatz von Triple P eine Verhaltensänderung von Kindern, die als Wirkung definiert wird. Um Triple P einsetzen zu dürfen, ist der Erwerb von Lizenzen und die Teilnahme an entsprechenden Kursen zur Ausbildung von Triple P Trainer*innen erforderlich.

Die Wirkungsmessung in diesem SIB erfolgt durch den Einsatz von drei verschiedenen Fragebögen, dem Strength and Difficulties Questionnaire (SDQ), dem Erziehungsfragebogen (EFB) und dem Client Satisfaction Questionnaire (CSQ) (vgl. Hornung 2017: 4).

Der SDQ ist ein Screening-Verfahren zur Erfassung von Verhaltensauffälligkeiten und -stärken von Kindern und Jugendlichen und findet Anwendung als Eingangsdiagnoseverfahren in der Kinder- und Jugendpsychiatrie³.

Der Erziehungsfragebogen (EFB) ist die deutsche Version von „Parenting Scale“ nach Arnold, O’Leary, Wolff & Acker aus dem Jahr 1993, einem Fragebogen der elterliches Erziehungsverhalten bei problematischen Verhaltensweisen von Kindern ermitteln soll (vgl. Feldmann 2006: 46).

Der CSQ wird regelhaft von Triple P als Evaluationsinstrument eingesetzt, zur Auswertung und Beurteilung der angebotenen Kurse.

Es handelt sich bei Triple P zwar um ein „Elterstraining“, es geht aber tendenziell von einem problematischen kindlichen Verhalten aus. Eltern werden dazu angeleitet, mit zweifelhaften Erziehungsmethoden die als problematisch definierten kindlichen Verhaltensweisen, hierzu zählen dann u.A. auch Schlafprobleme, zu bearbeiten. Die eingesetzten Erziehungsmethoden sind derart, dass Schäden an der Psyche von Kindern durch Verstörung, Beunruhigung oder Verängstigung nicht auszuschließen sind und einer gesunden Eltern-Kind-Beziehung entgegenstehen können. Der Einsatz von Triple P in den Hilfen zur Erziehung kann also als nicht unproblematisch beschrieben werden (zur Kritik an Triple P vgl. Deegener, Hurrelmann 2002; Gogol 2019: 132ff). Letztlich erscheint der Einsatz von Triple P in diesem SIB aufgrund der schnellen und messbaren Erfolgsaussichten und nicht aufgrund pädagogischer Abwägungen erfolgt zu sein. Kurzfristige Erfolge werden hier auch zum Preis unbekannter möglicher psychischer Spätfolgen erzielt.

Investitionen in welche Wirkung?

Wie oben bereits beschrieben, soll die Mehrzahl der Social Impact Bonds Arbeitsfähigkeit (wieder-)herstellen. Als gesellschaftliche Wirkung wird vornehmlich auf zukünftige Kosteneinsparungen für den Sozialstaat abgezielt. Auch wenn in Deutschland der Fokus in der Argumentation für SIBs nicht auf der finanziellen Entlastung des Sozialstaats liegt, wird dies immer wieder auch mit angeführt. SIBs entstanden aus der Idee, Kosten für den Sozialstaat langfristig einzusparen und dabei privaten Investor*innen finanzielle Renditen zu verschaffen. Es geht also im Kern um ein Investment mit möglichst hohen Renditeerwartungen. Um hohe Renditen zu legitimieren, wird das hohe Risiko der Investition betont. Dieses Risiko wird in SIBs durch das mögliche Verfehlen von zuvor festgelegten Wirkungszielen der sozialen Interventionen definiert. Wirkungsziele, die soziale Renditen ermöglichen sollen, orientieren sich beim Social Impact Investing, zu dem auch SIBs zählen, an monetär erfassbaren Renditen. Sie fokussieren auf zukünftige Einsparungen und/oder auf zukünftige Steuereinnahmen. Instrumente wie der Social Return on Investment versuchen eine monetäre Renditeerzielung sozialer Wirkungen abbilden zu können. Die Mehrzahl der SIBs streben die zu erreichenden Kosteneinsparungen durch die Aktivierung von Humankapital, der (Re-) Integration in den Arbeitsmarkt an. Dies wird entweder unmittelbar durch

³ Der SDQ wurde zur Abklärung dissozialer, oppositioneller, hyperaktiver und emotionaler Störungen von Kindern und Jugendlichen entwickelt. Er existiert in drei Versionen, einem Fragebogen für Eltern, einem für Lehrer*innen und einer Selbstbeurteilung für Kinder, was eine Kontrastierung der einzelnen Ergebnisse ermöglicht und so beispielsweise mögliche Verzerrungseffekte von Selbst- und Fremdwahrnehmung erkannt und eingeordnet werden können (vgl. Lohbeck et al. 2015: 233) Welche Versionen in diesem Social Impact Bond zum Einsatz kommen sollen, ist bislang nicht bekannt.

Arbeitsvermittlung oder mittelbar durch die Wiederherstellung bzw. Aufrechterhaltung von Gesundheit, Beendigung oder Vermeidung von Obdachlosigkeit oder einer Steigerung des Bildungsniveaus versucht. Auch in den Bereichen der Kinder- und Jugendhilfe und der Straffälligenhilfe können die angestrebten Einsparungen durch zukünftige bessere Vermittelbarkeit auf dem Arbeitsmarkt zusätzlich hochgerechnet werden.

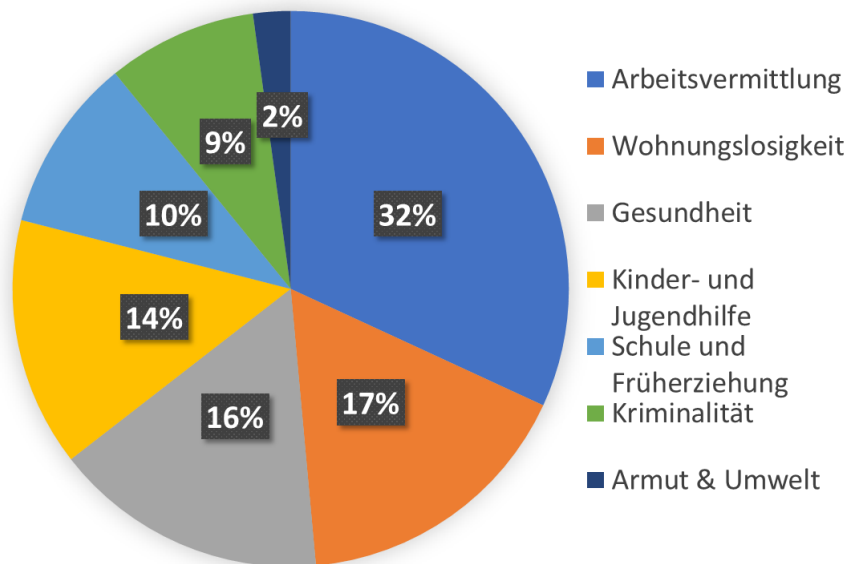


Abbildung 3: Verteilung SIBs nach Bereichen (eigene Darstellung bezogen auf die bis 2021 durchgeführten und laufenden SIBs⁴)

Investitionen in Humankapital sollen an die Stelle von Daseinsvorsorge und der Kompensation von sozialen Problemlagen treten. Es kommt infolge dessen zu einer Verschiebung der Verantwortlichkeiten auf das Individuum (vgl. Wohlfahrt 2007: 3).

Soziale Dienstleistungen müssen dann zukünftig durch den Nachweis von Wirksamkeit ihre Finanzierung, die sozialstaatlichen Investitionen, rechtfertigen. SIBs werden dabei zu einem sozialpolitischen Instrument, das geeignet erscheint, Denkmuster dahingehend zu ändern und Wirkungsorientierung im Sinne von Renditenerzielung (sozial und monetär), also von einem ökonomischen Standpunkt aus, zu betreiben. Kostenreduktion soll dadurch erreicht werden, dass anders als bei der auf die innere Verwaltung abzielende Effizienzsteigerung der Neuen Steuerung, angestrebte Wirkungen zu langfristigen Kostensenkungen führen. Hier kommt es zu einer Verknüpfung von ökonomischen und sozialen Zielen. Als soziale Wirkung gilt dann, was sich kostensenkend auswirkt und wird als effizientes, als wirtschaftliches Handeln Sozialer Dienste und als ein Ausdruck der Qualität der erbrachten Leistungen definiert. Wirksamkeit wird so auch zu einer Bedingung für das wirtschaftliche Erbringen von Leistungen, „Da eine unwirksame Leistung nicht wirtschaftlich sein kann [...]“ (Bundesregierung 2016: 312). Über eine solche Definition von Wirtschaftlichkeit wird es auch möglich, die Finanzierung sozialer Maßnahmen zu legitimieren, die einen höheren Ressourceneinsatz bedürfen, dafür aber die vereinbarten Wirkungen (also die vorab definierten Ziele) erreichen (vgl. Burmester 2020: 375f). Zielerreichung wird als

⁴ laut Social Finance UK (2021).

Qualitätsmerkmal dann höher eingestuft als die Prozessqualität, wie in den beiden oben dargestellten SIBs erkenntlich wurde.

In einer solchen Wirkungslogik wird nicht mehr nach den Ursachen für soziale Notlagen und den daraus resultierenden Hilfebedarfen gefragt. Strukturelle Armut, Arbeitslosigkeit, prekäres Wohnumfeld etc. sind lediglich Faktoren, die zu defizitären und unerwünschten Verhaltensweisen führen, die zu verändern sind (vgl. Burmester, Wohlfahrt 2015: 569).

Ob die prognostizierten zukünftigen Einsparungen tatsächlich erfolgen, kann heute noch nicht ermittelt werden. Darüber hinaus werden bei den Prognosen wichtige Faktoren, wie etwa die wirtschaftliche Entwicklung, die Entwicklung von Arbeitslosenzahlen und von Löhnen und Gehältern außer Acht gelassen. Zudem sind bis zum Eintreten der erhofften Einsparungen aufgrund weiterer, teils noch unbekannter Faktoren (wie etwa das aktuelle Pandemiegesehen), diese Einsparungen dann als Wirkung der Maßnahme nicht mehr nachvollziehbar. Dies ist im Konstrukt von SIBs allerdings auch nicht erforderlich, es wird durch das Erreichen bestimmter Ziele auch das Erreichen zukünftiger Wirkungen postuliert.

Fazit

Der Versuch einer Wirkungsmessung, wie etwa durch die Errechnung eines *Social Return on Investment*⁵ oder mittels der *Theory of Change*⁶ bleibt sehr aufwändig und erscheint nur schwer objektivierbar, was den Ansprüchen einer wissenschaftlich fundierten Messung von Kausalzusammenhängen im Sinne einer evidenzbasierten Sozialen Arbeit nur schwer Rechnung tragen kann. „Veränderungen können aber auch durch einen oder mehrere Störfaktoren entstehen, der bzw. die zu Verzerrungen („bias“) führen – dann ist interne Gültigkeit nicht mehr gegeben und die gefundenen Wirkungen sind nicht mehr eindeutig der Maßnahme zuzuordnen.“ (BMZ 2008: 15). Es bleibt daher oftmals bei Vermutungen und Annahmen über Kausalzusammenhänge, insbesondere da es kaum möglich erscheint, alle relevanten Störfaktoren zu identifizieren und in die Messung mit einzubeziehen. Darüber hinaus erscheint es nicht möglich, generelle Indikatoren für bestimmte soziale Problemlagen zu bilden. So kann etwa der Grad der Integration nur durch eine Vielzahl an Indikatoren abgeleitet werden, die sich auf verschiedene Teilaspekte beziehen, wie etwa Kenntnisse der Landessprache, Mitgliedschaft in Vereinen, etc. Die Kosten einer Wirkungsmessung in Relation zu dem Nutzen werden so schnell ausufernd und nicht mehr plausibel begründbar (vgl. Kettiger, Schwander 2011: 126f).

⁵ Der Social Return on Investment -Ansatz versucht aus verschiedenen Blickwinkeln die sozialen Wirkungen der Investition einer Sozialleistung zu beschreiben, in Geldwerte umzurechnen und damit die getätigten Investitionen zu legitimieren. Die zugrunde liegenden normativen Vorstellungen des Konzepts gehen davon aus, einen Mehrwert von Sozialausgaben, die hier als Investitionen angesehen werden, monetär bestimmen zu können. Der Ansatz erscheint allerdings wenig geeignet valide Wirkungsaussagen treffen zu können (hierzu ausführlicher vgl. Wohlfahrt, Ziegler 2019).

⁶Die Wirkungsmessung in der Theory of Change erfolgt auf Basis einer als Wirkungskette dargestellten Prozessbeschreibung, eines systematischer Projektplanungsprozess. Es findet hier also eher eine interne Evaluation der Maßnahme statt, da nicht versucht wird, die Wirkung anhand von (randomisierten oder rekonstruierten) Kontrollgruppen zu messen, sondern lediglich ermittelt wird, ob die mit Hilfe der Theory of Change entwickelten Interventionen in diesem speziellen, konstruierten Kausalzusammenhang funktionieren (vgl. Repp 2013: 31f).

Dass, ausgehend von der europäischen Sozialpolitik, die Idee der Wirkungsmessung als Steuerungsinstrument sozialer Dienste trotz der großen Schwierigkeiten und unüberwindbaren Hindernissen einer Verallgemeinerung der Ergebnisse, insbesondere im Hinblick auf eine Reproduzierbarkeit von Wirkungen, weiter voran getrieben wird, gründet nicht (mehr) nur in der Legitimierung von sozialen Dienstleistungen, sondern gerät zunehmend zur Forderung, nur noch solche sozialen Dienstleistungen zu finanzieren, die nachweislich wirksam sind. Der Sozialstaat geriete andernfalls in die Gefahr (auch weiterhin) die knappen Ressourcen fehlerhaft einzusetzen (vgl. Burmester, Wohlfahrt 2018b: 57). Als sinnvolle Investitionen erscheinen dann solche, die eine möglichst weitreichende Wirkung entfalten. Einzelfallhilfe wird so sukzessive zu einer Kostenfalle ohne ausreichende Renditeversprechen, die durch eine Finanzierung von „Hilfen zur Selbsthilfe“, sozialräumlichen Interventionen und einer Stärkung der Bürgergesellschaft, einer Subventionierung des freiwilligen Engagements begegnet werden soll. Die Frage nach der Wirkung von Maßnahmen bezieht sich dann also nicht mehr auf einen individuellen, von den beteiligten Fachkräften durchgeführten Reflexionsprozess bezogen auf den Einzelfall, um diesen mit Hilfe wissenschaftlicher Erkenntnisse deuten und die Frage der individuellen Wirksamkeit einer bestimmten Maßnahme einordnen zu können. Vielmehr bezieht sich die Wirkungsorientierung in diesem Kontext auf die Forderung nach einer konkreten Handlungsanweisung für bestimmte, trennscharfe Problemkonstellationen (vgl. Otto et al. 2010: 11).

Der Sozialstaat positioniert sich als Gesamtmanager der Sozialen Arbeit und nimmt als solcher direkten Einfluss auf Methoden und somit auch auf die Durchführung Sozialer Arbeit. Dies erreicht er, indem Wirtschaftlichkeit, Effektivität und Effizienz zu Primärzielen der Sozialen Arbeit erhoben werden. Im sozialstaatlichen Fokus stehen damit weniger die Ursachen komplexer sozialer Problemlagen oder die Lebenswelten der Betroffenen, die sich ebenso komplex darstellen, sondern die Effizienz der verrichteten Arbeit der einzelnen Sozialarbeiter*innen. „Diese Entwicklung ist Ausdruck dafür, dass Wirkungsforschung sich heute zunehmend weniger mit der Überprüfung des Erreichens gesellschaftspolitisch gesetzter Ziele beschäftigt (also z.B. Armutsbekämpfung, Schaffung von Ausbildungsplätzen), sondern vorrangig organisationale Leistungsmessung mit Effizienzfokus ist.“ (Dahme, Wohlfahrt 2010: 209f).

Das Eingrenzen von Methoden und Handlungsoptionen, die Festlegung auf wirkungsorientierte Maßnahmen und die Primärziele Effektivität, Effizienz und Wirtschaftlichkeit führen letztlich zu einem Verdrängungsprozess von Fachlichkeit und professionellem Handeln. So verspricht die Ausrichtung an einer outcomeorientierten Finanzierung auch eine Orientierung an einer evidenzbasierten Sozialen Arbeit. Dieser wird unterstellt, einfache Antworten auf komplexe Probleme geben zu können, die richtigen „Rezepte“ zur Bearbeitung sozialer Problemlagen zu liefern und letztlich die Kosten von Maßnahmen genau kalkulieren zu können. So soll sie es möglich machen, sogenannte *practice guidelines* zu erstellen, die Maßnahmen im Hinblick auf die angestrebte Zielerreichung in einer solchen Art und Weise kleinschrittig zu beschreiben und mit konkreten Handlungsanweisungen zu versehen, dass ein erfolgreicher Abschluss nahezu garantiert werden kann. Ein modularer Aufbau der methodisch-praktischen Vorgehensweise ermöglicht, den Fortschritt der Zielerreichung genau einzuschätzen und zu identifizieren, wann der nächste Schritt, das nächste Modul, zu bearbeiten ist (vgl. Dahmen 2011: 88f). Die Bereitstellung modularer Maßnahmen, die eine wirkungsorientierte Zielerreichung ermöglichen sollen, führt nicht nur zu einer Eingrenzung und Festlegung von Maßnahmen, sondern auch zu einer „Sozialen

Arbeit" die von Nicht-Fachkräften mindestens genauso gut ausgeführt werden kann. Es könnte sogar unterstellt werden, Nicht-Fachkräfte sind insofern besser geeignet, als dass sie aufgrund ihrer fehlenden Expertise die methodisch-praktische Umsetzung der Handlungsanweisungen weniger hinterfragen und so den zu erwartenden Erfolg der Maßnahme weniger gefährden. Der so ermöglichte Einsatz von angelernten Kräften in Bereichen, welche früher in der sozialen Dienstleistungsproduktion von Fachkräften besetzt werden mussten, ermöglicht den Sozialunternehmen einen neuen Zugriff auf den größten Kostenfaktor des Unternehmens, die produzierende Fachkraft. Es kann also erwartet werden, dass Sozialarbeiter*innen zukünftig vermehrt andere Aufgaben zugewiesen bekommen und/oder mit einer weiteren Absenkung der Reallöhne zu rechnen haben.

Social Impact Bonds erscheinen als ein sozialpolitisches Werkzeug zum nachhaltigen und fortschreitenden Umbau des fürsorgenden Sozialstaats, der sich verantwortlich für die Kompensation sozialer Notlagen fühlt, hin zu einem aktivierenden Sozialstaat neoliberaler Prägung, der als workfare-regime die Produktion von Humankapital als oberste Priorität ansieht. Dieses sozialpolitische Werkzeug soll gleichsam dazu genutzt werden, unmittelbaren Einfluss auf die Soziale Arbeit zu nehmen und diese entsprechend der neuen Prioritäten auszurichten. Dem in Bezug auf die sozialgesetzlich verankerte Wirkungsorientierung zu begegnen, erfordert den Einsatz von Wirkmodellen, die nicht von ökonomischen Gesichtspunkten geprägt sind und einen sozialwissenschaftlichen Wirkungsbegriff zugrunde legen, der nicht versucht ist, soziale Wirkungen zu monetarisieren und so monetäre Wirkungen auf den Einsatz sozialer Interventionen zurückzuführen.

Sebastian Gogol M.A. Studium der Sozialen Arbeit und Master-Studium „Management in sozialwirtschaftlichen und diakonischen Organisationen“. Arbeitet als Dipl.-Sozialarbeiter / Dipl.-Sozialpädagoge in der ambulanten Eingliederungshilfe der Diakonie Ruhr Wohnen gemeinnützige GmbH. Nebenberufliche Tätigkeit als wissenschaftlicher Mitarbeiter in verschiedenen Projekten zum Thema Wirksamkeit in der Sozialen Arbeit. Lehrauftrag zur Ökonomisierung des Sozialbereiches an der Ev. Fachhochschule Rheinland-Westfalen-Lippe.

Literatur

- BMZ. 2008: Wirkungsevaluierungen. Zum Stand der internationalen Diskussion und dessen Relevanz für die Evaluierung der deutschen Entwicklungszusammenarbeit.
- Bundesregierung. 2016: Regierungsentwurf. Entwurf eines Gesetzes zur Stärkung der Teilhabe und Selbstbestimmung von Menschen mit Behinderungen (Bundesteilhabegesetz – BTHG).
- Burmester, M. 2020: Wirksamkeit. Ein neuer ökonomischer Beurteilungsmaßstab für soziale Dienstleistungen. In H.-U. Otto (Hg.), Soziale Arbeit im Kapitalismus. Gesellschaftliche Verortungen – professionelle Positionen – politische Herausforderungen. Weinheim, Basel: Beltz Juventa, 366–379.
- Burmester, M., Friedemann, J. 2020: Soziale Dienstleistungen als Investition? Ökonomische Aspekte in der aktuellen Wirkungsdebatte. In M. Burmester, J. Friedemann, S. C. Funk, S. Kühnert, D. Zisenis (Hg.), Die Wirkungsdebatte in der Quartiersarbeit. Wiesbaden: Springer Fachmedien Wiesbaden; Imprint Springer VS, 197–210.
- Burmester, M., Wohlfahrt, N. 2015: Der soziale Wirkungskredit – die Lösung sozialer Probleme durch wirkungsorientiertes Investieren? Neue Praxis, 45. Jg., Heft 6, 559–571.

- Burmester, M., Wohlfahrt, N. 2018a: Der Social Impact Bond. Konzept und Implementierung. Düsseldorf.
- Burmester, M., Wohlfahrt, N. 2018b: Wozu die Wirkung sozialer Arbeit messen? Eine Spurensicherung. Berlin: Deutscher Verein für Öffentliche und Private Fürsorge e.V.; Lambertus.
- Cohen, R. 2014: Revolutionising Philanthropy: Impact Investment. Handout following speech at Mansion House on 23.1.14, letzter Aufruf 02. Juli 2018.
- Dahme, H.-J., Wohlfahrt, N. 2010: Evidenzbasierte Soziale Arbeit und wettbewerblich gesteuerte Sozialwirtschaft. In H.-U. Otto, A. Polutta, H. Ziegler (Hg.), What Works - Welches Wissen braucht die soziale Arbeit? Zum Konzept evidenzbasierter Praxis. Opladen: Budrich, 203–216.
- Dahmen, S. 2011: Evidenzbasierte soziale Arbeit? Zur Rolle wissenschaftlichen Wissens für sozialarbeiterisches Handeln. Baltmannsweiler: Schneider Hohengehren.
- Deegener, G., Hurrelmann, K. 2002: Kritische Stellungnahme zum Triple P. Positive Parenting Program Positives Erziehungsprogramm, https://kinderschutzbund-bayern.de/wp-content/uploads/kritische_stellungnahme_triplep_Prof_Deegener.pdf, letzter Aufruf 17. Juli 2018.
- Europäische Kommission. 2013: Mitteilung Der Kommission an das Europäische Parlament, den Rat, den europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen. Sozialinvestitionen für Wachstum und sozialen Zusammenhalt - einschließlich Durchführung des Europäischen Sozialfonds 2014–2020, [http://www.europarl.europa.eu/registre/docs_autres_institutions/commission_europeenne/com/2013/0083/COM_COM\(2013\)0083_DE.pdf](http://www.europarl.europa.eu/registre/docs_autres_institutions/commission_europeenne/com/2013/0083/COM_COM(2013)0083_DE.pdf), letzter Aufruf 13. Juli 2018.
- Feldmann, M. 2006: Bibliothherapie: Evaluation des Elternratgebers Triple P. Dissertation. Braunschweig.
- Fliegau, M. T. 2014: Sozialer Wirkungskredit. Struktur, Potenziale, Herausforderungen., https://www.stiftung-nv.de/sites/default/files/impulse_sozialer_wirkungskredit.pdf, letzter Aufruf 13. Juli 2018.
- GO Lab. 2021: Impact Bond Dataset, <https://golab.bsg.ox.ac.uk/knowledge-bank/indigo/impact-bond-dataset-v2/>, letzter Aufruf 30. März 2021.
- Gogol, S. 2019: Soziale Innovationen für eine wirksamere Praxis? Auseinandersetzung mit Social Impact Bonds im Landkreis Osnabrück und in Mannheim. Neue Praxis, 49. Jg., Heft 2, 125–139.
- Hornung, M. 2017: Prävention in den Hilfen zur Erziehung stärken im Landkreis Osnabrück.
- Hornung, M. 2018: Bildungschancen für Kinder verbessern in der Stadt Mannheim.
- Kettiger, D., Schwander, M. 2011: Wirkungsorientierung in der Sozialen Arbeit. Möglichkeiten und Grenzen. In A. Fritze, B. Maelicke, B. Uebelhart (Hg.), Management und Systementwicklung in der sozialen Arbeit. Baden-Baden: Nomos Verl.-Ges, 114–133.
- Lohbeck, A., Schultheiß, J., Petermann, F., Petermann, U. 2015: Die deutsche Selbstbeurteilungsversion des Strengths and Difficulties Questionnaire (SDQ-Deu-S). Diagnostica, 61. Jg., Heft 4, 222–235.
- NAB. 2014: Wirkungsorientiertes Investieren: Neue Finanzierungsquellen zur Lösung gesellschaftlicher Herausforderungen, <http://www.bertelsmann->

- stiftung.de/fileadmin/files/Projekte/59_Social_Investment/NAB-Abschlussbericht_Wirkungsvolles_Investieren.pdf, letzter Aufruf 30. Juni 2018.
- Otto, H.-U., Polutta, A., Ziegler, H. 2010: Zum Diskurs um evidenzbasierte Soziale Arbeit. In H.-U. Otto, A. Polutta, H. Ziegler (Hg.), What Works - Welches Wissen braucht die soziale Arbeit? Zum Konzept evidenzbasierter Praxis. Opladen: Budrich, 7–28.
- Petrick, S., Weber, M. 2013: Was sind Social Impact Bonds?: Definition, Strukturen, Marktentwicklung. Wirkungsorientierte Finanzierung für gesellschaftliche Herausforderungen, http://www.impactinmotion.com/wp/wp-content/uploads/2013/09/Machbarkeitsstudie_SIB_Auszug_20131007_Final.pdf, letzter Aufruf 30. Juni 2018.
- Repp, L. 2013: Soziale Wirkungsmessung im Social Entrepreneurship. Herausforderungen und Probleme. Wiesbaden: Springer Fachmedien Wiesbaden.
- Scheck, B., Weber, M. 2012: Impact Investing in Deutschland. Bestandsaufnahme und Handlungsanweisungen zur Weiterentwicklung, http://impactinmotion.com/wp-content/uploads/2013/05/Impact-Investing-in-Deutschland_FINAL.pdf, letzter Aufruf 02. Juli 2018.
- Social Finance Global Network. o.J.: Global Network - Social Finance, <http://socialfinance.org/global-network/>, letzter Aufruf 25. März 2021.
- Social Finance UK. 2021: Impact Bond Global Database, <https://sibdatabase.socialfinance.org.uk/>, letzter Aufruf 31. März 2021.
- Stadt Mannheim. o.J.: Leitfaden für das Mannheimer Unterstützungssystem Schule (MAUS). Förderzeitraum 2012/2013 und 2013/2014, https://www.mannheim.de/sites/default/files/page/2665/leitfaden_maus_iii.pdf, letzter Aufruf 30. März 2021.
- Triple P Deutschland GmbH. o.J.: Triple P auf einen Blick - Triple P, <http://www.triplep.de/de-de/triple-p-auf-einen-blick/>, letzter Aufruf 17. Juli 2018.
- Wohlfahrt, A., Ziegler, H. 2019: Zahlen und Geschichten - SROI als Umwertungstechnik im Sozialsektor. Neue Praxis, 49. Jg., Heft 2, 109–124.
- Wohlfahrt, N. 2007: Ausgrenzung durch Inklusionspolitiken. Prekariat. Ethik und Gesellschaft, Heft 1.

Barbara Flatters & Susanne Giel

Soziale Arbeit wirkt – und es gibt Wege, dies zu zeigen. Erfahrungen aus Evaluationen

Die Ausgangssituation

Soziale Arbeit ist heutzutage grundsätzlich an Wirkungen ausgerichtet. In der Definition Sozialer Arbeit ist enthalten, dass sie „gesellschaftliche Veränderungen, soziale Entwicklungen und den sozialen Zusammenhalt sowie die Stärkung der Autonomie und Selbstbestimmung von Menschen“ fördert (Deutscher Berufsverband für Soziale Arbeit e.V. 2016).

Unstrittig ist, dass soziale Arbeit wirkungsorientiert sein soll, die Herausforderung besteht jedoch in der Mess- und Belegbarkeit dieser Wirkung. Das hat verschiedene Gründe: Soziale Arbeit wirkt in der Regel nicht linear und monokausal. Vielmehr fußen die Erfolge Sozialer Arbeit auf einem gut abgestimmten Bündel von Maßnahmen und sind vielfältigen Einflüssen ausgesetzt. Zudem ist zu bedenken, dass Soziale Arbeit nicht standardisiert erfolgt, sondern bedarfsorientiert und situationsadäquat umgesetzt wird.

Um Wirkungen einer Intervention¹ nachzuweisen, ist häufig das erste, was einem einfällt, ein Experiment durchzuführen. Dies ist uns insbesondere aus dem medizinischen Bereich geläufig, wenn beispielsweise die Wirksamkeit eines neuen Medikaments in verblindeten, randomisierten kontrollierten Studien nachgewiesen werden soll. Der Einsatz solcher Designs ist aber im Bereich der Sozialen Arbeit aus verschiedenen Gründen meist nicht möglich: So finden Interventionen Sozialer Arbeit nicht im „luftleeren Raum“ statt – es wäre kaum möglich, die Bedingungen zwischen Kontroll- und Interventionsgruppe gleichzuhalten. Allein sich der Bedingungen bewusst zu sein, ist vor Einsatz der Intervention herausfordernd. Eine Aufteilung in Kontroll- und Interventionsgruppe kann darüber hinaus auch aus ethischer Sicht schwierig sein. Außerdem birgt der Einsatz experimenteller Designs im Bereich der Sozialen Arbeit nur eingeschränkte Nützlichkeit: Hätte man in einem Experiment die Wirksamkeit einer Intervention nachgewiesen, ist eine Wiederholung oder Übertragung der Intervention nicht möglich, wenn man nicht weiß, wie die Wirkung zustande kam. Oder – noch ungünstiger – bei einer Intervention, die sich als wirkungslos erweist, weiß man nicht, an welchen Stellschrauben man zur Verbesserung ansetzen kann.

Diese Umstände erfordern schon seit Jahren eine hohe Kreativität im empirischen Nachweis von Wirkungen. Evaluator*innen stehen regelmäßig vor der Herausforderung, im Feld der Sozialen Arbeit Wirkungen zu belegen, und haben dazu eine Palette an Lösungen entwickelt: Um der Komplexität Sozialer Arbeit gerecht zu werden und um praxisrelevantes Wissen zu erzeugen, bieten sich vor allem sogenannte theoriebasierte Evaluationsansätze an. Hierbei werden die Grundannahmen der Beteiligten darüber, wie Soziale Arbeit das Angestrebte bewirken soll, zur Basis der Evaluation gemacht (Giel 2013). Diese Annahmen werden zunächst in einer

¹ Der Begriff „Intervention“ wird in diesem Beitrag als Sammelbegriff für jede Art von Projekt, Programm, Maßnahme, Initiative usw. verwendet. Damit kann beispielsweise ein Beratungsangebot oder die alltägliche Arbeit in einer Einrichtung gemeint sein. Wir rufen die Leser*innen dazu auf, hier jeweils an ihr eigenes Arbeitsfeld zu denken.

Programmtheorie² expliziert. Diese Programmtheorie wird dann im weiteren Verlauf der Evaluation empirisch überprüft. Auch die im Folgenden dargestellte Kontributionsanalyse gehört zu den theoriebasierten Evaluationsansätzen.

Die Kontributionsanalyse

Relativ bekannt ist die Nutzung von Wirkungsmodellen zur Konstruktion und empirischen Überprüfung von Wirkungszusammenhängen (z. B. Funnell und Rogers 2011). Ebenfalls prominent ist die „Realistic Evaluation“ nach Pawson und Tilley (2010), die die Frage in den Mittelpunkt rückt: „Was wirkt für wen unter welchen Bedingungen?“. Einen in jüngster Zeit viel beachteten Vorschlag zur Beantwortung dieser Fragestellung bietet die „Kontributionsanalyse“ nach John Mayne (2012).

Bei der Kontributionsanalyse will man herausfinden, ob die Intervention „einen Unterschied gemacht hat“. Dabei wird von einem etwas ungewöhnlichen Kausalitätsverständnis ausgegangen, dem der „Kontribution“. Spricht man von Kausalität, kann man Ursachen dahingehend unterscheiden, ob sie notwendig und/oder hinreichend für ein Ergebnis sind (Mayne 2012: 275, eigene Beispiele):

- Das Absolvieren einer Berufsausbildung im dualen System ist eine hinreichende, aber nicht notwendige Ursache für das Erlernen eines Berufs – wer eine Berufsausbildung meistert, ist danach auf seinen Beruf vorbereitet, es gibt aber auch andere Wege, einen Beruf zu erlernen, beispielsweise ein Studium.
- Ein Wohngeldanspruch ist eine notwendige, aber nicht hinreichende Ursache, um Wohngeld zu erhalten: Wer diesen Anspruch nicht hat, kann kein Wohngeld erhalten, der Anspruch allein genügt aber nicht, um Wohngeld zu bekommen (ist also nicht hinreichend), denn darüber hinaus muss die betreffende Person noch einen Antrag auf Wohngeld stellen.
- Notwendig und hinreichend sind Ursachen für ein Ergebnis, wenn nur bei Vorliegen dieser Ursache das Ergebnis ausgelöst werden kann (notwendig) und es gleichzeitig ausreicht, dass diese Ursache vorliegt, um das Ergebnis auszulösen (hinreichend). Hier fällt es schwer, ein geeignetes Beispiel zu finden, was verdeutlicht, wie selten solche Gegebenheiten sind.

Häufig sind Ursachen aber weder notwendig noch hinreichend für ein Ergebnis, haben aber trotzdem einen Einfluss darauf: Wenn Eltern ihre Kinder bei den Hausaufgaben unterstützen können, ist dies förderlich für deren schulischen Erfolg. Die Unterstützung ist aber keine notwendige Bedingung: Auch Kinder, die nicht von ihren Eltern unterstützt werden, können in der Schule erfolgreich sein. Gleichzeitig ist die Unterstützung aber auch keine hinreichende Bedingung, beispielsweise ist darüber hinaus notwendig, dass die Lehrer*innen guten Unterricht machen und das Kind sich konzentrieren kann.

Sachverhalte, bei denen Ursachen weder notwendig noch hinreichend sind, aber trotzdem einen Beitrag zu einem Ergebnis leisten, kommen in unserer Gesellschaft ständig vor. Auch bei

² Auch hier gibt es verschiedene Begrifflichkeiten, auf deren Unterschiede hier nicht eingegangen werden kann. Daher wird der Begriff „Programmtheorie“ hier als Sammelbegriff für Wirkmodelle, Wirkungsmodelle, Theory of Change usw. verwendet. Zentral ist, dass die Wirkungsannahmen einer Intervention explizit erfasst und ggf. schematisch dargestellt werden.

Interventionen der Sozialen Arbeit können wir weder davon ausgehen, dass sie notwendig für bestimmte angestrebte Wirkungen sind (dass also das Erreichen ohne die Intervention unmöglich wäre) noch, dass sie hinreichend sind (also allein und unabhängig von Rahmenbedingungen zur gewünschten Wirkung führen). Trotzdem leisten Interventionen in der Sozialen Arbeit einen wichtigen Beitrag zu angestrebten Wirkungen: Im Zusammenspiel mit anderen Ursachen (z.B. weiteren Interventionen und Rahmenbedingungen) sind Interventionen erfolgreich und dieser Mehrwert soll in Evaluationen nachgewiesen werden³. Man will herausfinden, ob die Intervention „einen Unterschied gemacht hat“. Hier setzt die Kontributionsanalyse an.

Ziel der Kontributionsanalyse ist es, den Beitrag, einer Intervention zu einem gewünschten Ergebnis so genau wie möglich beschreiben zu können. Die Unsicherheit über den Beitrag einer Intervention zu einem gewünschten Ergebnis soll reduziert werden. Dazu soll ein besseres Verständnis dessen, weshalb bestimmte Ergebnisse auftreten, generiert werden. Hierfür werden „alternative Erklärungen“ betrachtet und „unterstützende Einflussfaktoren“ mit einbezogen. Eine Kontributionsanalyse soll die Frage beantworten: Hat die Intervention in Anbetracht der zahlreichen Faktoren, die ein Ergebnis beeinflussen, einen bemerkenswerten Beitrag zum beobachteten Ergebnis geleistet und – wenn dem so ist – auf welche Art und Weise? (Mayne 2012:273).

Die Wirksamkeit einer Intervention wird in der Kontributionsanalyse dann angenommen, wenn Folgendes erfüllt ist (Mayne 2012: 272):

1. Der Intervention liegt eine gut begründete Programmtheorie zugrunde: Die Wirkungspfade und zugrunde liegenden Annahmen sind plausibel, stichhaltig, durch bestehende Forschung untermauert und von Stakeholdern unterstützt.
2. Die Aktivitäten der Intervention wurden wie geplant durchgeführt (d. h. wie in der Programmtheorie angenommen).
3. Die Programmtheorie wurde durch empirische Evidenz verifiziert: Die Ergebnisse sind wie erwartet eingetreten (d. h., dass auch „Zwischenschritte“ wie in der Programmtheorie erwartet durchlaufen wurden).
4. Externe Faktoren, die die Intervention und ihren Erfolg beeinflussen könnten, wurden untersucht und es wurde festgestellt, dass sie nicht relevant für das Ergebnis sind, oder – wenn sie relevant sind – wurden sie in die Einschätzung des Beitrags der Intervention einbezogen.

Das Vorgehen einer Kontributionsanalyse lässt sich in sechs Schritten beschreiben (vgl. Abbildung 1):

³ Bei Zusammenhängen, bei denen man von notwendigen und/oder hinreichenden Ursachen ausgeht, ist es möglich, diese „kontrafaktisch“ nachzuweisen: Will man belegen, dass eine Ursache notwendig für ein Ergebnis ist, kann man versuchen Fälle als „Gegenbeweis“ zu finden, bei denen das Ergebnis vorliegt, ohne dass die als notwendig angenommene Ursache vorhanden ist – ist dies nicht möglich, hat man den kontrafaktischen Beweis erbracht. Ebenso bei als hinreichend angenommenen Ursachen: Fälle, bei denen die als hinreichend angenommene Ursache vorliegt, ohne, dass das Ergebnis wie erwartet ist, wären hier Gegenbeispiele. Kann man diese nicht finden, hat man hier den kontrafaktischen Beweis erbracht. Bei Ursachen, die weder als hinreichend noch als notwendig angenommen werden, kann nicht kontrafaktisch argumentiert werden.

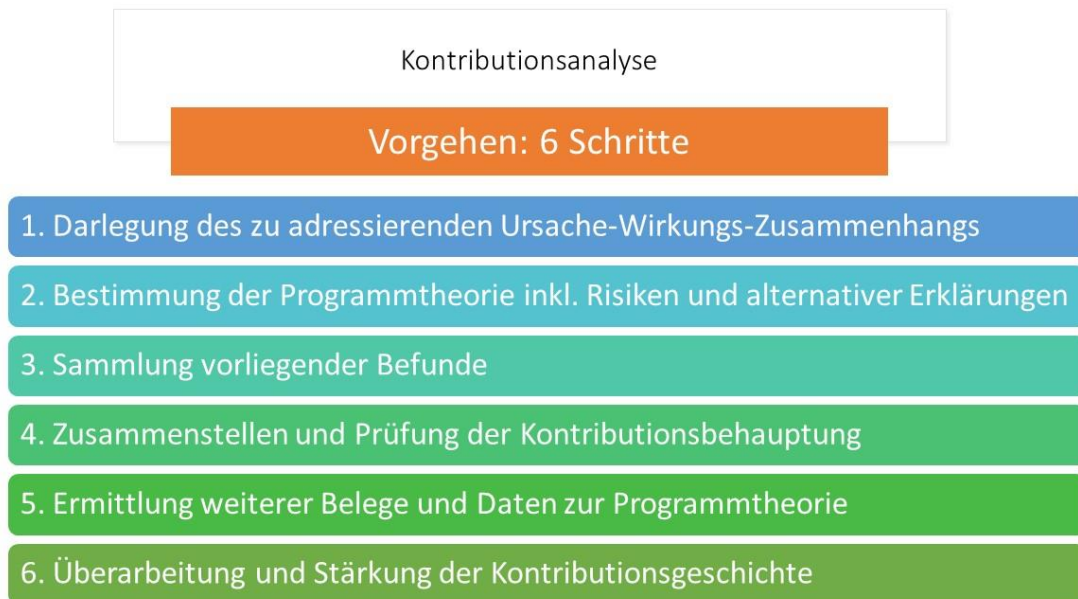


Abbildung 1: Die sechs Schritte der Kontributionsanalyse nach John Mayne (2012, S. 272, eigene Übersetzung)

Im ersten Schritt⁴ wird entschieden und festgehalten, welcher Ursache-Wirkungs-Zusammenhang betrachtet werden soll. Üblicherweise findet dies in Form einer Frage nach dem Muster „Inwiefern hat die Intervention X zum Ergebnis Y beigetragen?“ statt. Im zweiten Schritt wird die Programmtheorie expliziert, also herausgearbeitet, auf welchem Weg und mit welchen „Zwischenschritten“ die Intervention zum Ergebnis beitragen soll. Bereits an dieser Stelle wird überlegt, welche Bedingungen relevant sein könnten und welche alternativen Erklärungen es jeweils für das Auftreten der Zwischenergebnisse geben könnte. Hier ist eine schematische Darstellung der Programmtheorie hilfreich, um tatsächlich alle Zwischenschritte im Blick zu behalten (Mayne 2015). Im dritten Schritt werden bereits bestehende Belege – sowohl für die einzelnen Schritte der Programmtheorie als auch für die „alternativen Erklärungen“ – gesucht. Hier können beispielsweise Meta-Studien oder Datenbanken, in denen Interventionen und dazugehörige Evidenzen zusammengetragen werden, hilfreich sein. Im vierten Schritt werden diese Belege zusammengefasst, eine „Kontributionsbehauptung“ wird entworfen und es wird geprüft, welche Unklarheiten in Bezug auf den Beitrag der Intervention zum Ergebnis noch bestehen. Im fünften Schritt findet dann empirische Arbeit statt, um die im vierten Schritt festgestellten „Lücken“ zu schließen. Im sechsten Schritt werden die verschiedenen empirischen Befunde genutzt, um die „Kontributionsgeschichte“ weiter anzureichern und plausibel zu machen. Daraufhin wird das Vorgehen ab Schritt vier wiederholt: die „Kontributionsbehauptung“ wird in Hinblick auf mögliche Erklärungslücken geprüft, es finden weitere empirische Erhebungen statt, um diese Lücken zu schließen und die neuen und alten Erkenntnisse werden in einer angereicherten „Kontributionsgeschichte“ erzählt. Die Anzahl der Wiederholungsschleifen ist in der Evaluationspraxis allerdings begrenzt.

⁴ Eine detaillierte Beschreibung des Vorgehens in diesen sechs Schritten anhand von mehreren Beispielen findet sich in Delahais und Toulemonde 2012.

Das Vorgehen der Kontributionsanalyse am Beispiel „Qualitätssiegel Lebensort Vielfalt“

Das Vorgehen einer Kontributionsanalyse wird im Folgenden an einem Beispiel aus unserer Evaluationspraxis erläutert. Beim gewählten Beispiel handelt es sich um eine Evaluation, die sich noch in der Umsetzung befindet. Es geht um die Evaluation des Projekts „Qualitätssiegel Lebensort Vielfalt“ der Schwulenberatung Berlin (SBB)⁵. Die Zielstellung des Projekts besteht darin, dass ambulante und stationäre Pflegeeinrichtungen sowie Hospize einen Prozess durchlaufen, an dessen Ende die Berücksichtigung der Belange von LSBTI⁶-Personen steht, was nach dem Durchlaufen eines Prüfverfahrens mit der Vergabe eines Qualitätssiegels belegt wird. Die SBB geht auf die Einrichtungen zu, berät und begleitet Leitungspersonen und Qualitätsverantwortliche, organisiert und realisiert Schulungen für die Mitarbeitenden der Einrichtungen und befähigt die Einrichtungen so, die notwendigen Kriterien für die Siegelvergabe zu erfüllen.

Finanziert wird das Projekt durch den Verband der Privaten Krankenversicherung, kurz PKV-Verband, der auch die Evaluation beauftragt hat. Die Evaluation soll verschiedene Funktionen erfüllen. Sie soll die Wirkungsfähigkeit belegen und vorhandenen Optimierungsbedarf identifizieren. Ebenfalls charakterisierend für diese Evaluation ist es, dass sie explorativ angelegt ist, d.h.: Zunächst soll der modellhafte Prozess optimiert und sein Wirkungspotential umfassend identifiziert werden. Wenn sich im Rahmen des aktuellen Durchlaufs das intendierte Wirkungspotential zeigt, dann soll der Ansatz skaliert und verbreitet werden.

Schritt 1 – Den zentralen Wirkungszusammenhang identifizieren

Entsprechend der Kontributionsanalyse galt es zunächst festzustellen, was eigentlich der zentrale Ursache-Wirkungs-Zusammenhang ist, der verfolgt wird und den es zu belegen gilt. Dazu gehört eine gründliche Beschäftigung mit dem Evaluationsgegenstand, seinem Vorgehen und den damit verfolgten Zielen. In diesem Beispiel ist bemerkenswert, dass sich in den Gesprächen mit den Beteiligten (mit dem Projektteam der SBB und dem PKV-Verband) deutlich abzeichnete, dass die Prozesse nicht nur darauf abzielten, dass die Belange von LSBTI-Personen in der Pflege besser berücksichtigt werden. Darüberhinausgehend konnte festgestellt und belegt werden, dass das Projekt einen Beitrag zur Gesundheitsförderung leisten kann. Die Gesundheitsförderung umfasst dabei alle Bewohnenden und die Mitarbeitenden der Pflegeeinrichtungen.

Schritt 2 – Programmtheorie bestimmen

Nachdem der zentrale, mit dem Projekt verfolgte Wirkungszusammenhang geklärt war, galt es, die dem Projekt unterliegende Programmtheorie zu bestimmen. Wir begaben uns auf die Suche nach Wirkungsannahmen, die dem Projekt zugrunde liegen und auf die dahinter liegende Wirkungslogik. Dazu analysierten wir Projektdokumente (Konzepte, Berichte, Diversity Check Katalog etc.), führten eine Fokusgruppe mit dem Projektteam durch und stimmten daraus abgeleitete Ergebnisse mit dem PKV-Verband und dem Projekt ab.

⁵ Weitere Informationen zum Projekt finden sich hier: <https://schwulenberatungberlin.de/qualitaetssiegel-lebensort-vielfalt/>

⁶ LSBTI steht für Lesbisch, Schwule, Bisexuelle, Transsexuelle, Intersexuelle

Eine zentrale Frage in diesem Arbeitsschritt ist immer, wie sich diese Wirkungsannahmen darstellen lassen. Wir haben uns dem Wirkungsmodell des Programmbaums bedient (Bartsch u.a. 2016). Der Programmbaum ist ein Logisches Modell, das ein Raster dafür bietet, die Planung oder auch Herstellung von Wirkungen von Programmen, Projekten oder einzelnen Maßnahmen darstellen und beschreiben zu können. Der Programmbaum besteht aus Wurzeln und einem Grund, auf dem ein Programm oder Projekt entwickelt und umgesetzt wird, den Bedingungen. Ausgehend von den Wurzeln besteht der Stamm aus einem Konzept und Aktivitäten: Zunächst wird ein Konzept vor dem Hintergrund der Bedingungen entwickelt, das dann im nächsten Schritt in Form von Aktivitäten umgesetzt wird. Aus diesen Aktivitäten erwachsen Resultate wie Outputs, hier dargestellt als Äste, an diesen wachsen Früchte, einerseits Outcomes – Veränderungen auf der Ebene von Individuen – und Impacts – strukturelle Veränderungen.⁷

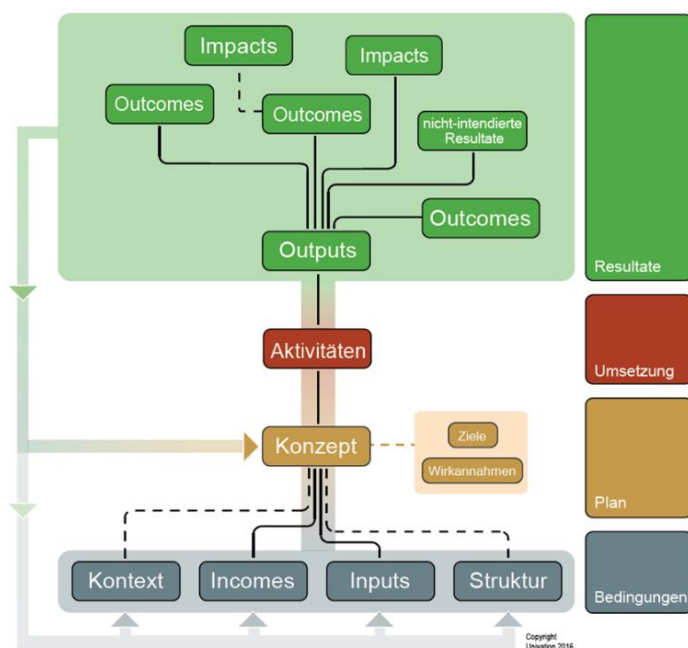


Abbildung 2: Der Programmbaum[©] von Univation

Für das Projekt Qualitätssiegel Lebensort Vielfalt ist das untenstehende, mit konkreten Inhalten gefüllte Wirkungsmodell entstanden (vgl. Abbildung 3). Da es sich um ein Modell handelt, also nicht um eine naturnahe wirklichkeitsgetreue Abbildung, müssen Vereinfachungen vorgenommen werden. Abgebildet sind die für wichtig erachteten Bedingungen, die zentralen Aktivitäten und Resultate, sowohl Outcomes⁸ als auch Impacts⁹.

⁷ Eine kurze Erläuterung des Programmbaums findet sich in folgendem Video: <https://www.youtube.com/watch?v=7yNVpSHJmSc>

⁸ Outcomes beschreiben Veränderungen bzw. Stabilisierungen auf individueller Ebene im Wissen, bei Fähigkeiten oder Fertigkeiten, in den Einstellungen, bei Kompetenzen oder sozialen Werten: [https://eval-wiki.org/glossar/Outcomes_\(eines_Programms\)](https://eval-wiki.org/glossar/Outcomes_(eines_Programms))

⁹ Impacts umfassen Veränderungen und Stabilisierungen auf struktureller Ebene: [https://eval-wiki.org/glossar/Impacts_\(eines_Programms\)](https://eval-wiki.org/glossar/Impacts_(eines_Programms))

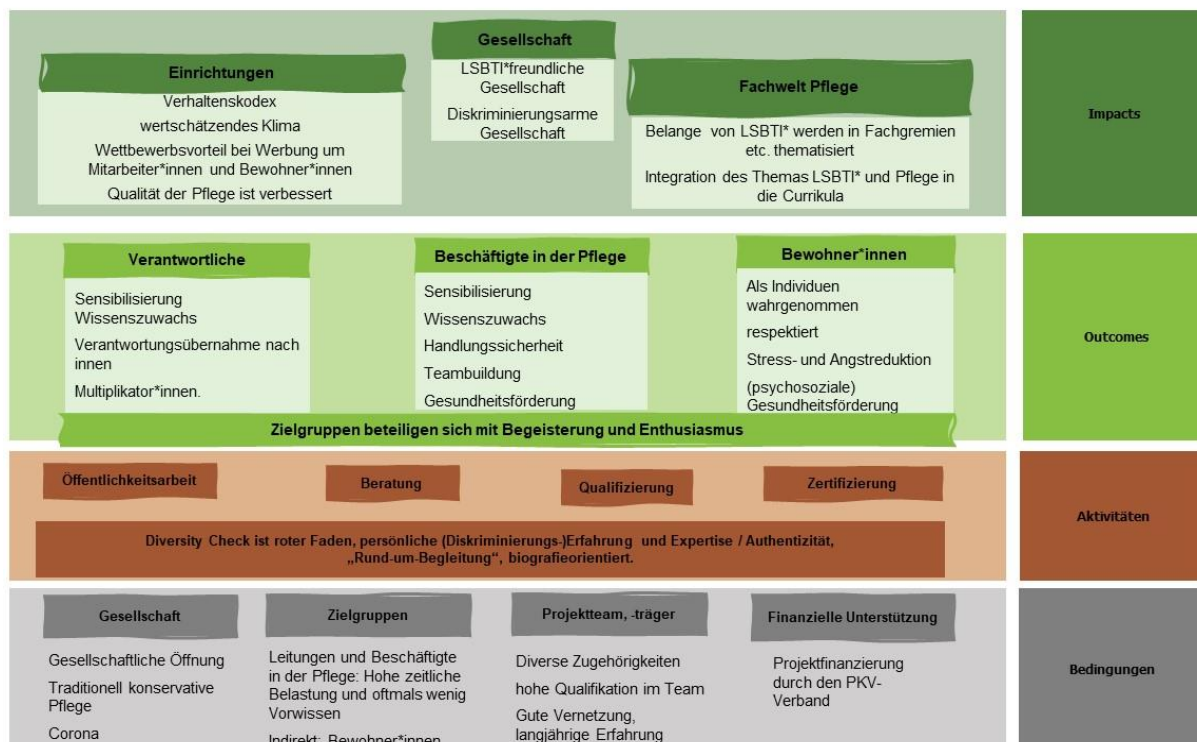


Abbildung 3: Das vorläufige Wirkungsmodell des Qualitätssiegels Lebensort Vielfalt

Beispielhaft sollen im Folgenden einige der herausgearbeiteten Annahmen über Wirkungspfade dargestellt werden.

Zunächst ist festzuhalten, dass es sich beim „Qualitätssiegel Lebensort Vielfalt“ um ein sogenanntes Kaskadenprojekt¹⁰ handelt, d. h.: Zunächst werden die Leitungen der Pflegeeinrichtungen eingebunden. Leitungspersonen sollen den gesamten Entwicklungsprozess für die Organisation rahmen. Damit werden Voraussetzungen für die Umsetzung von strukturellen Veränderungen geschaffen, wie z. B. Anpassungen im Leitbild, Veränderungen im Personalmanagement, in der Neudefinition von Angehörigen etc. Demzufolge besteht eine zentrale Wirkungsannahme des Projekts darin, dass nicht nur Individuen qualifiziert werden müssen, sondern auch die Rahmenbedingungen in den Einrichtungen verändert und insgesamt das Klima verbessert werden müssen, um die Einrichtungen zu Lebensorten der Vielfalt werden zu lassen. Dies drückt sich beispielsweise in Hausordnungen und Besuchsregeln aus und wird durch einen entwickelten Verhaltenskodex belegt.

Ausgehend von einer ersten Prüfung wird der Fortbildungsbedarf für die Mitarbeitenden der Einrichtung bestimmt und entsprechende Qualifizierungen umgesetzt. Dabei gibt es sogenannte Pflichtmodule, wie beispielsweise die Pflege von LSBTI, zu den Themen HIV, Intersexualität und Transsexualität in der Pflege. Als Outcomeziele sind formuliert: Mitarbeitende sind sensibilisiert für die Situation und Bedarfe von LSBTI, sie wissen, was bei der Pflege von LSBTI zu berücksichtigen ist und erlangen die Kompetenzen das zu tun. Letztlich sollen die Bewohner*innen bzw. Klient*innen das sensible und kompetente Handeln der Mitarbeitenden erleben, und damit stressfreier und gesünder leben.

¹⁰ <https://eval-wiki.org/glossar/Kaskadenprogramm>

Die Bedingungen sind ebenfalls in die Programmtheorien einzubeziehen: Exemplarisch sei die Annahme des Projektteams benannt, dass eine wichtige Bedingung für das Gelingen des Vorgehens ist, dass das Projektteam selbst LSBTI-divers zusammengesetzt ist. Hierin sieht das Projektteam eine wichtige Voraussetzung dafür, dass es seine Zielgruppen nicht nur kognitiv, sondern auch emotional erreicht.

Schritt 3 – Sammlung vorliegender Befunde

Nach der Bestimmung der inneren Logik und der Wirkungsannahmen aus Projektsicht und aus Sicht des Projektfinanziers beginnt die Suche nach Studien und Belegen für die angenommenen Wirkungszusammenhänge. Zur Illustration seien hier drei Beispiele genannt.

Ein erster Schritt bestand in Recherchen zum Zusammenhang zwischen sexueller, geschlechtlicher Identität und Gesundheit. Hier ist die Datenlage aus unterschiedlichen Gründen eher dünn. Tatsächlich gibt es aber bereits eine erste Metaanalyse aus Deutschland, die einen guten Überblick über die Gesundheitsrisiken von LSBTI gibt (Pöge u. a.). Dabei ist die zentrale Erkenntnis: Nicht LSBTI-Sein ist das Gesundheitsrisiko, sondern die (gesellschaftlichen) Bedingungen stellen für LSBTI-Personen ein Gesundheitsrisiko dar, denn LSBTI-Personen sind vielfältigen unterschiedlichen Diskriminierungserfahrungen ausgesetzt. Und tatsächlich, der Zusammenhang zwischen Diskriminierung und Gesundheit ist vielfältig belegt (siehe z. B. Igel u. a. 2010). Dadurch lässt sich eine zentrale Strategie des Projekts untermauern: Es wirkt darauf hin, dass Diskriminierung in der Pflege abgebaut wird. Beispielsweise zielen viele der geprüften Kriterien exakt auf den Abbau von Diskriminierung.

Die Grundüberlegung des Projekts, dass das Lernen (beispielsweise in den Fortbildungen für die Mitarbeitenden) die Lernenden auch emotional einbeziehen soll, findet zahlreiche Bestätigung in der Forschung (Arnold/Holzapfel 2008). Dass nicht zuletzt auch in der Erwachsenenbildung Emotionen, Motivationen und Kognition miteinander zu integrieren sind, ist unbestritten.

Für eine weitere zentrale Wirkungsannahme des Projekts, dass die Veränderungen in den Pflegeeinrichtungen als Organisationsentwicklungsprozess angelegt sein müssen, finden sich ebenfalls einige zentrale Hinweise in der Literatur. Zu nennen sind das Drei-Phasen-Modell nach Kurt Lewin¹¹, ebenso das Konzept der Lebensweltorientierung und der Settingansatz,¹² der auch im Leitfaden Prävention eine zentrale Rolle spielt (GKV-Spitzenverband 2018). Zusammenfassend berücksichtigt die Herangehensweise zentrale, bereits bestätigte Annahmen, wie beispielsweise, dass

- eine Veränderung auf der Organisationsebene ein langfristiger Prozess ist, der eine kontinuierliche Begleitung erfordert,
- nicht nur Menschen qualifiziert, sondern auch Rahmenbedingungen verändert werden müssen,
- Veränderungen nachhaltig verankert werden müssen, indem beispielsweise ein Rezertifizierungsprozess für die Einrichtungen vorgesehen ist.

¹¹ Kurt Lewin (z. B. 1947) entwirft ein Modell, das davon ausgeht, dass zunächst die Beteiligten aufgetaut (unfreezing) werden müssen, um im nächsten Schritt bewegt werden zu können (changing). Damit die Veränderung nachhaltig bestehen bleibt, müssen die Veränderungen in neue Strukturen gegossen werden (freezing).

¹² Der Settingansatz geht davon aus, dass Gesundheitsförderung in den Sozialzusammenhängen, in denen sich Menschen in ihrem Alltag aufhalten, zu betreiben ist

Schritt 4 – Zusammenstellung und Prüfung der Kontributionsbehauptungen

Die eingesammelten Informationen werden nun zusammengetragen und miteinander abgeglichen. Das heißt, für die in Schritt 2 postulierten Wirkungsannahmen werden externe Belege oder Widersprüche recherchiert (Schritt 3) und vergleichend betrachtet.

Zusätzlich wird festgehalten, welche weiteren Einflüsse in diesem Geflecht wirken bzw. wirken könnten. Beispielsweise ist davon auszugehen, dass sich die Qualität der Pflege (unabhängig von der Teilnahme am Projekt „Qualitätssiegel Lebensort Vielfalt“) auf die Gesundheit von Mitarbeitenden und Bewohner*innen bzw. Klient*innen auswirkt. Möglicherweise beteiligen sich an dem Projekt solche Einrichtungen, die ohnehin einen hohen Qualitätsanspruch haben. Umgekehrt ist denkbar, dass Einrichtungen Qualitätsmängel mit dem Gütesiegel Lebensort Vielfalt kaschieren möchten. In beiden Fällen nimmt die Qualität Einfluss auf die Gesundheit von Bewohner*innen und Klient*innen.

Naheliegender ist daneben auch die Frage danach, welche Einrichtungen sich überhaupt an dem Vorhaben beteiligen. Sind es vielleicht nur solche, die ohnehin für LSBTI aufgeschlossen sind? Kontributionsanalytisch gesprochen heißt das: Leistet eine LSBTI-sensible Leitung eine notwendige oder hinreichende Ursache für die Wirksamkeit des Qualitätssiegels?

Ein Zwischenbericht, der unterschiedliche Einflussfaktoren benennt, in Frage stellt und bestätigt, hat im Sinne einer Kontributionsanalyse ebenfalls die Aufgabe festzuhalten, welche Aspekte der Programmtheorie empirisch noch nicht ausgeleuchtet sind. Die ersten Analysen zeigen zwar, dass das Projekt im Großen und Ganzen das Potential aufweist, intendierte Wirkungen zu erbringen. Allerdings ist noch nicht bestätigt, inwieweit dieses Konzept auch in der konkreten Umsetzung aufgeht und ob möglicherweise noch weitere unerwartete Wirkungen auftreten.

Schritt 5 – Ermittlung weiterer Beleg und Daten zur Programmtheorie

Die Ermittlung weiterer Belege ist die Phase, in der sich die Evaluation aktuell (im Frühjahr 2021) befindet. Um zu beantworten, inwieweit es gelingt, die postulierten Wirkungsannahmen in der Praxis zu entfalten, ist es im vorliegenden Beispiel zentral, die Perspektiven der beteiligten Stakeholder einzubeziehen. Neben dem Projektteam war von Beginn an die Einbeziehung von Einrichtungsleitungen und Mitarbeitenden geplant. Ein Teil der Leitungspersonen wurde im Rahmen einer digitalen Fokusgruppe, ein anderer Teil im Rahmen von telefonischen Einzelinterviews nach ihrer Motivation zur Teilnahme, nach ihrem Erleben des Prozesses und nach wahrgenommenen Veränderungen auf der Ebene von Bewohner*innen/Klient*innen, Mitarbeitenden und der Einrichtung als Ganzem befragt. Eine Gruppe von Mitarbeitenden der Einrichtungen wird im Rahmen einer Fokusgruppe befragt, für alle anderen Mitarbeitenden ist eine (weitgehend standardisierte) Online-Befragung vorgesehen.

Eine weitere wichtige Informationsquelle werden die Prüfberichte der Zertifizierungsverfahren sein. Hierin finden sich Hinweise auf die vorgenommenen Änderungen, wie auch Ergebnisse von Befragungen von Mitarbeitenden, von Bewohner*innen und Klient*innen.

In einem nächsten Schritt, allerdings nach Abschluss der aktuell laufenden nur siebenmonatigen Evaluation, sollten Klient*innen und Bewohner*innen der Einrichtungen befragt werden. Denkbar sind hier exemplarische Gruppenerhebungen wie auch schriftliche Befragungen.

Schritt 6 – Überarbeitung und Stärkung der Kontributionsgeschichte

An dieser Stelle ist ein abschließender Bericht, die Kontributionsgeschichte, zu schreiben und zu erzählen: Hierbei sind

- Belege für die Wirkungszusammenhänge zusammenzutragen,
- voraussichtlich auch eine Modifizierung der Programmtheorie und des Wirkungsmodells vorzunehmen,
- die befördernden und die hindernden Bedingungen auszuführen.

Da unser abschließender Bericht endet, bevor die Bewohner*innen einbezogen wurden, ist dann auch festzuhalten, dass diese Prüfung noch aussteht und für die Zukunft empfohlen wird.

Chancen der kontributionsanalytischen Herangehensweise

Die besondere Chance der kontributionsanalytischen Herangehensweise an die Evaluation im hier beschriebenen Beispiel besteht darin, dass die Evaluation bereits in einer Entwicklungs- und Modellphase des Projekts auf Wirkungen ausgerichtet werden kann. Die Frage nach Wirkungen kann auch ohne experimentelle Vergleichsgruppen in sehr umfänglicher Art und Weise, unter Einbeziehung verschiedenster Perspektiven (Beteiligte, Betroffene, Forschung und Wissenschaft) untersucht werden. Sie ist in der Lage, sehr komplexe Projekte und Vorhaben zu untersuchen, in denen auf unterschiedlichen Ebenen Ursache-Wirkungs-Gefüge greifen.

Eine weitere Stärke ist das schrittweise (und möglicherweise auch iterative) Vorgehen der Kontributionsanalyse: Die beschriebene Vorgehensweise ermöglicht die detaillierte Betrachtung der einzelnen Schritte einer Programmtheorie, so dass wenn nötig Stellschrauben zur Verbesserung identifiziert werden können. Dies ermöglicht, auch solche Vorhaben auf Wirkungen hin zu untersuchen, die sich noch in der Entwicklung befinden, sich den Bedingungen anpassen oder verschiedene Interventionen erproben müssen. Damit ermöglicht dieses Evaluationsdesign die Nutzung von Hinweisen auf Verbesserungspotentiale bereits im Verlauf der Evaluation.

Am Ende steht eine fundierte Darstellung bestätigter Wirkungszusammenhänge sowie die Benennung offen gebliebener Fragen. Nicht zuletzt möchten wir anmerken, dass aus unserer Sicht die Frage nach dem Beitrag, den Soziale Arbeit leisten kann, viel wirklichkeitsnäher ist als die Frage nach vollständiger oder spezifisch anteiliger Zurechnung von Wirkungen.

Barbara Flatters (M.A. Politikwissenschaft) arbeitet als Evaluatorin und Projektmanagerin bei der IMAP GmbH in Düsseldorf. Zuvor arbeitete sie bei Univation – Institut für Evaluation GmbH und an der Heinrich-Heine-Universität Düsseldorf. Für ihre Masterarbeit zum Thema „Relevanz von Evaluationen in deutschen Kommunen“ wurde sie mit dem Nachwuchspreis der Gesellschaft für Evaluation 2018 ausgezeichnet.

Dr. Susanne Giel (Dipl. Soziologin) ist seit 2006 als freie Evaluatorin tätig und seit 2014 Mitgesellschafterin von Univation – Institut für Evaluation GmbH. Sie evaluiert vor allem Programme und Projekte im Bereich Bildung, Kinder- und Jugendhilfe, Rassismusprävention und Gesundheitsförderung. Mit dem Anliegen, Wirkungen besser in den Blick nehmen zu können, promovierte sie zu theoriebasierten Evaluationsansätzen und deren methodischer Umsetzung.

Literatur

Arnold, Rolf; Holzapfel, Günther (Hrsg.) (2008): Emotionen und Lernen. Die vergessenen Gefühle in der (Erwachsenen-)Pädagogik. Baltmannsweiler: Gabler

- Bartsch, Samera; Beywl, Wolfgang; Niestroj, Melanie (2016): Der Programmbaum als Evaluationsinstrument, in: Susanne Giel; Katharina Klockgether und Susanne Mäder (Hrsg.): Evaluationspraxis. Professionalisierung – Ansätze – Methoden. 2., korrigierte und ergänzte Auflage. Münster, New York: Waxmann, S. 89–112.
- Delahais, Thomas, and Jacques Toulemonde. "Applying contribution analysis: Lessons from five years of practice." *Evaluation* 18.3 (2012): 281–293.
- Deutscher Berufsverband für Soziale Arbeit e.V. (2016): Deutschsprachige Definition Sozialer Arbeit. Verfügbar unter: <https://www.dbsh.de/profession/definition-der-sozialen-arbeit/deutsche-fassung.html>; zuletzt abgerufen am 14.12.2020.
- Funnell, Sue C.; Rogers, Patricia J. (2011): Purposeful program theory. Effective use of theories of change and logic models. 1. ed. San Francisco, CA: Jossey-Bass (Research Methods for the Social Sciences, v.31).
- GKV-Spitzenverband (2020): Leitfaden Prävention – Handlungsfelder in stationären Pflegeeinrichtungen nach § 5 SGB XI, verfügbar unter: https://www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/praevention__selbsthilfe__beratung/praevention/praevention_leitfaden/Leitfaden_Praevention_2020_barrierefrei.pdf; zuletzt abgerufen am 29.04.2021.
- Giel, Susanne (2013): Theoriebasierte Evaluation. Konzepte und methodische Umsetzungen. Münster.
- Igel, Ulrike; Brähler, Elmar; Grande, Gesine (2010): Der Einfluss von Diskriminierungserfahrungen auf die Gesundheit von MigrantInnen. In: *Psychiatrische Praxis* 37 (4), S. 183–190 und Morawa, Eva; Erim, Yesim (2014): Zusammenhang von wahrgenommener Diskriminierung mit Depressivität und gesundheitsbezogener Lebensqualität bei türkisch- und polnischstämmigen Migranten. In: *Psychiatrische Praxis* 41 (4), S. 200–207.
- Lewin, Kurt (1947): Frontiers in Group Dynamics: Concept, Method and Reality in Social Science; Social Equilibria and Social Change. In: *Human Relations* 1 (1), S. 5–41. Mayne, John. "Contribution analysis: Coming of age?" *Evaluation* 18.3 (2012): 270–280.
- Mayne, John (2012): Contribution analysis: Coming of age? In: *Evaluation* 18 (3), S. 270–280.
- Mayne, John (2015): Useful theory of change models. *Canadian Journal of Program Evaluation* 30.2 (2015).
- Pawson, Ray; Tilley, Nick (2010): Realistic evaluation. Reprinted. London: SAGE.
- Pöge, Kathleen; Dennert, Gabriele; Koppe, Uwe; Güldenring, Annette; Matthigack, Ev B.; Rommel, Alexander (2020): Die gesundheitliche Lage von lesbischen, schwulen, bisexuellen sowie trans- und intergeschlechtlichen Menschen. Online verfügbar unter https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsJ/JoHM_S1_2020_Gesundheitliche_Lage_LSBTI.pdf;jsessionid=EB833B5362FBBF641A13D7C564935BE2.internet082?__blob=publicationFile; zuletzt abgerufen am 29.04.2021.

Edgar Baumgartner, Aline Kaufmann & Michael Lambertus

Evaluationen in Organisationen der Sozialen Arbeit – Bedingungen und Faktoren der Angleichung und Variation von Evaluationsaktivitäten

1. Einführung

In der Sozialen Arbeit berichten verschiedene Beobachter und Beobachterinnen eine zunehmende Bedeutung von Evaluationen (vgl. Baier, Keller, Koch & Wigger, 2013; Dollinger, 2018; Petersén & Olsson, 2014; Widmer et al., 2016, S. 467). Der allgemeine Siegeszug der Evaluation (vgl. Stockmann & Meyer, 2020; von Werthern, 2019, S. 33) macht – so scheint es – auch vor dem Sozialbereich nicht Halt. Der vorliegende Beitrag möchte zunächst thematisieren, welche Entwicklungen diese Zunahme an Evaluationen begünstigen, und dann in einem weiteren Schritt empirische Hinweise zum Stand von Evaluationsaktivitäten in der Sozialen Arbeit im deutschsprachigen Raum geben (Abschnitt 1). Ein besonderer Fokus liegt auf der Frage, wie weit aus theoretischer Sicht eine Angleichung von Strukturen und Praktiken der Evaluation anzunehmen ist (Abschnitt 2), oder ob das unterschiedliche Vermögen sozialer Organisationen, Evaluationen (intern) durchführen zu können, eine diesbezügliche Variation schafft (Abschnitt 3). Der Beitrag gibt abschließend (Abschnitt 4) einen Ausblick auf ein laufendes Forschungsprojekt in der Schweiz zu Evaluationen in Organisationen der Sozialen Arbeit, das die Thematik Angleichung und Variation aufnimmt.

2. Entwicklungen und Stand von Evaluationen in der Sozialen Arbeit

Evaluationen von Interventionen und Programmen reichen in der Sozialen Arbeit weit zurück. Die Erforschung der Bedingungen, Formen, Ergebnisse und Wirkungen kann als inhärenter Bestandteil der Sozialen Arbeit betrachtet werden (vgl. Haubrich, 2008, S. 441). Der «Ruf nach Evaluation und Wirkungsforschung» (Schulze-Krüdener 2017, S. 16) ist entsprechend kein grundsätzlich neuartiger, an die Soziale Arbeit adressierter Anspruch. Vielmehr ist die Frage nach der Wirksamkeit mit der Idee von Professionalität untrennbar verbunden (vgl. Polutta, 2014), da das gesellschaftliche Mandat Sozialer Arbeit darauf fußt, dass die professionelle Problembearbeitung Wirkungen bzw. bei Adressaten und Adressatinnen einen Nutzen erzeugt (vgl. Baumgartner & Sommerfeld, 2022).

Seit den 1960er und 1970er Jahren werden Wirkungen Sozialer Arbeit systematisch untersucht (vgl. Dollinger, 2018, S. 246). Im Fokus stehen vor allem Modellvorhaben und entsprechend als «Begleitforschung» (vgl. Müller, 1978) angelegte Studien. Die Verbreitung von Evaluationen über Modellprojekte hinaus vollzieht sich im deutschsprachigen Raum mit der Transformation des Wohlfahrtsstaates und insbesondere mit der Durchsetzung der wirkungsorientierten Verwaltungsführung seit den 1990er Jahren. Da die Soziale Arbeit häufig in sozialstaatlichen Finanzierungs- und Steuerungsformen eingebettet ist (vgl. Becker-Lenz & Baumgartner, 2016), ist die Verwaltungsmodernisierung – etwa mit dem Modell des New-Public-Managements-, ein entscheidender Treiber der Institutionalisierung von Evaluation (vgl. Stockmann & Meyer, 2020). Aufgrund der Intention der öffentlichen Verwaltung, über Steuerung die Effektivität und Effizienz der Leistungserbringung zu erhöhen, wurden vor allem auf die Ziel- und Ergebnisreichung

ausgerichtete (summative) Evaluationen zu wichtigen Quellen steuerungsrelevanter Informationen (vgl. Hornbostel, 2012; Ziegler, 2012).

In der Sozialen Arbeit hat die formative Evaluation – eine auf die Optimierungen der Praxis angelegte Form der Evaluation (vgl. Stegmann & Schwab 2012, S. 12) – jedoch immer eine wichtige Rolle gespielt. In den 1990er Jahren konnte sich beispielsweise das Konzept der Selbstevaluation (vgl. Heiner, 1988) etablieren, das die «Reflexivität» professioneller Leistungserbringung auf eine (datengestützte) Basis stellt (vgl. Baumgartner & Sommerfeld, 2022): Fachkräfte selbst unterziehen ihr eigenes professionelles Handeln einer Evaluation und setzen hierfür auch spezifische Instrumente ein (vgl. Baumgartner, 2016, S. 243; von Spiegel, 2008, S. 277f.).

Aktuellere Impulse für organisationsinterne Evaluationsaktivitäten ergeben sich aus Konzepten wie «Wirkungsorientierung» und «evidenzbasierte Praxis». Das Konzept «evidenzbasierte Praxis» wurde von Gambrill (1999) in die Soziale Arbeit eingeführt und betont, dass moderne Professionalität im Kern auf gesichertem wissenschaftlichem Wissen (Evidenz) gründet (Sackett et al., 1996). Insbesondere das «bottom-up»-Verständnis «evidenzbasierter Praxis» (vgl. James, 2016) sieht die Möglichkeit vor, in Organisationen durch Evaluationen lokales Wissen über Erfolge und Wirkungen professionellen Handelns zu generieren, um die professionelle Selbststeuerung und die Optimierung der professionellen Leistungserbringung zu stärken. Das Konzept der «Wirkungsorientierung» (Polutta, 2014), das auf verschiedene Ebenen (Politik, Organisation, Programmen, professionelles Handeln) bezogen werden kann, stützt diese Intention, innerhalb von Organisationen der Sozialen Arbeit die Frage der Wirkungen eigener Leistungen und Angebote stärker zu gewichten und mit empirischen Mitteln zu bearbeiten.

Auch in Professionalitätskonzeptionen verankerte Entwicklungen haben somit – neben dem «Managerialismus» (Otto & Ziegler, 2011) – Impulse für vermehrte Evaluationsaktivitäten in der Sozialen Arbeit gesetzt (vgl. Merchel, 2015; Ziegler, 2012). Doch wie weit sich dies in der Verbreitung und Praxis von Evaluationen in der Sozialen Arbeit niederschlägt, lässt sich nur bruchstückhaft aufzeigen. Denn im deutschsprachigen Raum geben nur wenige Studien einen Überblick über die Evaluationspraxis. Für die Schweiz lassen sich drei Studien mit engem Bezug zur Wirkungsmessung anführen: Bieri und Studer (2016) berichten von einem Survey bei Non-Profit-Organisationen, die ein Zewo-Gütesiegel (Zertifizierung für gemeinnützige, Spenden sammelnde Organisationen) aufweisen (mit 223 Fällen, 2009): Sie kommen zum Ergebnis, dass 40% der Organisationen Wirkungsmessungen durchführen. Eine zweite, neuere Studie, welche bei Mitgliedorganisationen von drei Fachverbänden im Sozial- und Gesundheitsbereich durchgeführt wurde, berechnet, dass 54% der Organisationen die langfristige Zielerreichung, u.a. auf Basis von Selbstbeurteilungen, überprüfen (vgl. Bernet & Gmür, 2015). Die Autoren schätzen den Anteil der Organisationen, welche für die Wirkungsmessung systematisch wissenschaftliche (Erhebungs- und Auswertungs-) Methoden nutzen, auf 20%. Eine weitere Studie stellt eine Erhebung von 2019 dar und berücksichtigt Organisationen, die in der Sozialen Arbeit in der Deutschschweiz tätig sind (vgl. Hochuli Freund, Gebert, John & Sprenger, 2020). Bei insgesamt 180 Organisationen liegen Angaben zur Bedeutung bzw. zu Aktivitäten der Wirkungsmessung vor. Rund 70% der adressierten fachlichen Leitungspersonen erachten es für ihre Organisation als wichtig, die Wirkungen der eigenen Leistungen und Angebote zu kennen. Drei Viertel der Organisationen in der Deutschschweiz geben an, interne Aktivitäten zur Messung von Wirkungen vorzunehmen, und

dabei sowohl informelle wie auch formelle Formen einzuschließen. In einem Fünftel der Einrichtungen gab es in den letzten fünf Jahren zudem eine externe Studie zwecks Messung der Wirkungen (vgl. Hochuli Freund, Gebert, John & Sprenger, 2020).

Eine Studie von Phineo (2013) zur Verankerung der Wirkungsorientierung in Non-Profit-Organisationen informiert über die Praxis der Wirkungsmessung in Deutschland (Basis: 83 Organisationen). Mehr als die Hälfte der Befragten bestätigt, dass in ihrer Organisation eine quantitative und qualitative Erhebung zur Überprüfung der tatsächlich erreichten Erfolge stattfindet. Rund zwei Drittel (67%) erachten es als (eher) zutreffend, dass die eigene Einrichtung zielgerichtet Wirkungen von ausgewählten Projekten oder Programmen erfasst. Die Bandbreite der eingesetzten Verfahren zur Untersuchung und Messung von Wirkungen ist dabei groß, wobei die Analyse der Beschwerden von Zielgruppen bzw. Leistungsempfängenden sowie eine Selbstevaluation mittels Befragungsverfahren am häufigsten zur Anwendung kommen (vgl. Phineo, 2013, S. 25). Auch eine aktuelle deutsche Studie zur Praxis der Evaluation in der psychosozialen Beratungsarbeit (vgl. Froncek et. al, 2018) bestätigt eine relativ starke Verankerung von Evaluationsaktivitäten in der Praxis. Von 566 befragten professionellen Beratenden bejahen 60% ($N = 339$), dass Evaluationen zu ihrer Beratungsarbeit stattfinden. Die gewichtigsten Themen der Evaluation sind die «Zufriedenheit mit der Beratung allgemein» sowie die «Erreichung der Ziele, die zu Beginn der Beratung definiert wurden». Eine Differenzierung in Beratungsarbeit in institutionellen Kontexten versus in eigener Praxis zeigt Unterschiede bei der Wahl der Evaluationsmethoden: Schriftliche, selbst entwickelte Fragebögen werden in der institutionellen Beratungsarbeit mit einem Anteil von 62% häufiger eingesetzt als in eigener Praxis (47%), umgekehrt finden systematische Beobachtungen, Gespräche (Interviews) und Falldokumentationen (Sekundäranalysen) in der institutionellen Beratungsarbeit seltener Verwendung (vgl. Froncek et al, 2018, S. 258). Die Präferenz einzelner Methoden kann mit unterschiedlichen Zwecken der Evaluation in Verbindung gebracht werden, insbesondere mit der Herausforderung, die Beratungsarbeit gegenüber Dritten darstellen zu können.

Diese punktuellen Verweise mögen ausreichend belegen, dass sich Evaluationsaktivitäten als Teil der Praxis in der Sozialen Arbeit zunehmend etabliert, diese jedoch nicht vollständig durchdrungen haben und daher arbeitsfeld- und organisationspezifische Unterschiede auszumachen sind.

3 Angleichung von Evaluationsaktivitäten aus neoinstitutionalistischer Perspektive

Aus einer organisationstheoretischen Perspektive stellt sich die Frage, warum eine Angleichung in Bezug auf Evaluationsaktivitäten stattfindet und gleichwohl von Variationen auszugehen ist. Um dieser Frage nachzugehen, ist die neoinstitutionalistische Organisationstheorie (vgl. Walgenbach & Meyer, 2008) und ihr Beitrag zur Erklärung von Organisationsverhalten in institutionalisierten Umwelten besonders relevant. Im Rahmen dieses Ansatzes wird davon ausgegangen, dass die Umwelt von Sozialen Organisationen aus institutionalisierten Erwartungsstrukturen besteht, welche die Ausgestaltung von Organisationen nachhaltig prägen (vgl. Walgenbach & Meyer, 2008, S. 11). Die Entsprechung solcher Erwartungsstrukturen und Rationalitätsansprüche der organisationalen Umwelt ist aus neoinstitutionalistischer Perspektive Ausdruck des Ziels von Organisationen, ihre Legitimität in einem zunehmend konkurrenzierenden Umfeld zu sichern (vgl. Meyer & Rowan, 1977). In diesem Kontext heißt das: Praktiken und Strukturen von Evaluationen werden in Organisationen folglich nicht aus Gründen der Effizienz oder zur Optimierung und

Überprüfung der Praxis eingeführt, sondern sind vielmehr auf Erwartungen in der Organisationsumwelt ausgerichtet (vgl. Meyer & Rowan, 1977, S. 341).

Nach DiMaggio und Powell (1983) richten sich Organisationen in einem organisationalen Feld nach einem «Sinnsystem» aus und handeln dabei immer aufeinander bezogen. Das Konzept des organisationalen Feldes schließt die gesamte relevante Umwelt mit ein, in der Organisationen operieren. Dazu zählen beispielsweise Organisationen, die in gemeinsame Netzwerke eingebunden sind, miteinander kooperieren oder untereinander im Wettbewerb stehen. Auch Ämter, Behörden und soziale Bewegungen, die Einfluss auf die Struktur, das Verhalten und das Überleben von Organisationen nehmen können, lassen sich zurechnen (Walgenbach & Meyer, 2008, S. 34). Das Sinn- und Regelsystem beinhaltet formelle und informelle Regeln, Wert- und Normvorstellungen sowie gemeinsam geteilte Überzeugungen. Solche institutionalisierten Regelsysteme in organisationalen Feldern lassen sich nach Scott in drei Säulen gliedern: Die regulative, normative und kognitive Säule (siehe Tabelle 1). Hier geht man von drei grundlegenden institutionellen Mechanismen aus, die unten näher beschrieben werden, welche sowohl auf Strukturen als auch die Praxis von Evaluationen in Sozialen Organisationen Einfluss nehmen können

	Regulative Säule	Normative Säule	Kulturell-kognitive Säule
Grundlage für Befolgung	Zweckmäßigkeit	Soziale Verpflichtung/ Verantwortung	Selbstverständlichkeit, gemeinsam geteilte Auffassungen
Grundlage für Ordnung	Regeln und Vorschriften	Verbindliche Erwartungen/ Erwartungshaltungen	Konstitutive Konzepte und Schemata
Indikatoren	Gesetze, Regeln, Vorschriften, Sanktionen	Zertifikate, Zulassungen, Akkreditierungen	Gemeinsam geteilte Überzeugungen und Handlungslogiken
Grundlage für Legitimität	Rechtliche/gesetzliche Sanktionen	Moralische Kontrolle	Nachvollziehbarkeit, Anerkennung, kulturell getragen
Mechanismen	Zwang	Normen und Werte	Nachahmung (Mimesis)
Beispiele	Staatliche Regulierungen, Leistungsaufträge	Professionelle Standards (Berufsverbände/Hochschulen)	Strukturen, Praktiken, Managementmodelle

Tabelle 1: Dreisäulenmodell nach Scott (in Anlehnung an Falk, 2016, 106)

DiMaggio und Powell (1983) sprechen in diesem Zusammenhang von «institutionellem Isomorphismus» und beschreiben dabei verschiedene Angleichungsprozesse (Isomorphie) in organisationalen Feldern. Die Autoren stellen die Frage, «What makes organizations so similar?» (DiMaggio & Powell, 1983, S. 147), das heißt, welche Faktoren begünstigen organisationale Angleichungsprozesse? Aus diesem Blickwinkel führt die starke Umweltbezogenheit von institutionalisierten Organisationen dazu, dass sich die Organisationen innerhalb eines organisationalen Feldes immer ähnlicher werden (Mayrhofer, 2014, S. 286). Dabei findet der Prozess der Angleichung mehr oder weniger unbewusst und freiwillig statt. Organisationen übernehmen bestimmte strukturelle Elemente, Konzepte oder Praktiken, um sich Legitimität in ihrer Umwelt zu verschaffen, und werden einander durch die gegenseitige Imitation ähnlicher (vgl. Walgenbach & Meyer, 2008, S. 63ff.).

DiMaggio und Powell (1991, S. 67ff.) beschreiben diese Prozesse der Strukturangleichung mit drei unterschiedlichen Mechanismen zur Herstellung von institutioneller Isomorphie wie folgt:

Das Konzept des «Coercive Isomorphism» (DiMaggio & Powell, 1991, S. 67ff.) bezeichnet die Strukturangleichung von Organisationen durch Zwang (siehe regulative Säule, Tabelle 1), welche auf Grund gesetzlicher Regulierungen oder staatlicher Vorgaben in Organisationsstrukturen aufgenommen werden müssen. Es kann von formellem und informellem Druck ausgegangen werden, welcher sich in staatlichen Vorgaben oder in rechtlichen Regulierungen zeigt (Mayrhofer, 2014, S. 287; Trenkmann, 2017, S. 64). Bezüglich Evaluationsaktivitäten in der Sozialen Arbeit sind hier der Einfluss der wirkungsorientierten Verwaltungsführung und die mögliche Koppelung von spezifischen Evaluationspraxen an staatliche Subventionen oder vorgegebene Berichterstattungspraxen zu nennen.

«Normative Pressures» (DiMaggio/Powell, 1991, S. 70ff.) ist eine zweite Quelle für Angleichungsprozesse aufgrund von bestehenden Normen. Hier geht es beispielsweise um professionelle Standards, die über mehr oder weniger expliziten, normativen Druck durch Wissenschaftlerinnen und Wissenschaftler sowie Berufsverbände eingefordert werden (siehe normative Säule, Tabelle 1). Professionalisierung wird als ein kollektiver Kampf der Institutionsmitglieder verstanden, um die Bedingungen und Methoden ihrer Arbeit zu bestimmen. Diese verbreiteten Professionsstandards wiederum bilden dann für Organisationen einen Orientierungsrahmen, der gemäß diesem Ansatz zu bestimmten Evaluationspraktiken führt (Mayrhofer, 2014, S. 287; Trenkmann, 2017, S. 64). Dies kann sich namentlich in der Auf- oder Übernahme von Konzepten wie «Wirkungsorientierung» oder «evidenzbasierte Praxis» zeigen, da es die eigene Reputation verbessert, beispielsweise die Wirksamkeit der erbrachten Dienstleistungen auszuweisen.

Im Gegensatz zu «Coercive Isomorphism» und «Normative Pressures» unterliegen dem dritten Konzept «Mimetic Processes» (DiMaggio & Powell, 1991, S. 69f.) nicht Zwang oder normativer Druck, sondern Angleichungsprozesse durch Nachahmung (siehe kulturell-kognitive Säule, Tabelle 1). Darunter ist das bewusste oder unbewusste Imitieren von Strukturen, Praktiken oder Managementmodellen von Organisationen, die als besonders erfolgreich erachtet werden, gemeint. Hier geht es um die Orientierung an anderen Organisationen, die innerhalb des organisationalen Feldes eine zentrale Stellung besitzen. Diese Imitationen resultieren oft aus Unsicherheit oder fehlendem Wissen über neue Technologien oder Umweltbedingungen (vgl. Mayrhofer, 2014, S. 287; Trenkmann, 2017, S. 64).

4. Organisationsinterne Bedingungen und Einflussfaktoren auf Evaluationsaktivitäten

Diesen beschriebenen Mechanismen und Kräften, welche zu einer Angleichung der Evaluationspraxis in Organisationen der Sozialen Arbeit führen können, steht jedoch gegenüber, dass Organisationen zur Umsetzung adäquater Evaluationsaktivitäten eventuell nicht in der Lage sind. Denn neben externen, in der Organisationsumwelt angesiedelten Entwicklungen, üben auch verschiedene organisationsinterne Voraussetzungen und Bedingungen einen Einfluss auf die resultierende Praxis und die Strukturen der Evaluation aus.

Eine entscheidende Größe hierbei ist «evaluation capacity»; (Naccarella et al. 2007, S. 232), das heißt das unterschiedliche Vermögen von Organisationen, Evaluationen zur eigenen Praxis durchführen zu können (vgl. Hense & Widmer, 2013, S. 261; Naccarella et al., 2007, S. 232). Der Aufbau eines solchen Vermögens, das «evaluation capacity building» (ECB, Stockmann, 2018, S. 368), kann auf individueller, organisationaler oder systemischer Ebene angesiedelt sein (vgl.

Stockmann, 2018, S. 368). In Bezug auf die hier relevante Organisationsebene haben sich mehrere Faktoren als förderlich erwiesen: Eine Verankerung von Evaluationen in der Organisationskultur, Fähigkeiten und Erfahrungen von Mitarbeitenden in Bezug auf Evaluationen sowie ausreichende finanzielle und zeitliche Ressourcen (vgl. Baker, Stephens & Hitchcock, 2010, S. 969; Cousins, Goh, Elliott & Bourgeois, 2014, S. 15; Gervin et al. 2010; Norton et al. 2016, S. 4f.). In ihrem Literaturreview ergänzen Carman und Fredericks (2010) die bedeutsame Rolle der Leitung (siehe auch Alaimo, 2008; Preskill & Boyle, 2008). Bieri und Studer (2016, S. 269) betonen zudem, dass die Institutionalisierung von Wirkungsmessungen durch die Verknüpfung mit bzw. die Integration in Management- oder Führungssystemen gefördert wird (zur Integration von Evaluationsaktivitäten in den Arbeitsalltag, siehe auch Bestvater & Beywl, 2015). Die Studie von Phineo (2013, S. 18) unterstreicht, dass bei Non-Profit-Organisationen ausreichende personelle und finanzielle Ressourcen für die Messung von Wirkungen ebenso eine Herausforderung darstellen wie auch das fehlende oder unzureichende Know-how von Mitarbeitenden. Bei den erschwerenden Faktoren nennt diese Studie jedoch auch die Schwierigkeit, messbare und inhaltlich sinnvolle Indikatoren oder Kennzahlen für die Erfassung von Wirkungen zu entwickeln (vgl. Phineo, 2013, S. 18). Das aktuelle Bild der Evaluationspraxis ist damit auch Ausdruck jener methodologischen Herausforderungen, die sich bei Evaluationen in der Sozialen Arbeit stellen (vgl. Baumgartner, 2016; Carman & Fredericks, 2010; Dollinger 2018; Schneider, 2016): Hierzu zählen beispielsweise der Mangel an eindeutigen, konsensfähigen und messbaren Erfolgskriterien (vgl. Baumgartner, 2016; Heiner, 2001, S. 488; Schneider, 2016, S. 205) oder fehlende einfache Kausalitäten zwischen Intervention und Wirkung (vgl. Bieri & Studer, 2016, S. 269; von Spiegel, 2008, S. 276). Eine weitere Schwierigkeit betrifft den Nachweis von Wirkungen, namentlich den Einsatz von hierfür adäquaten Designs. Denn in der Sozialen Arbeit ist die Verwendung eines randomisierten Kontrollgruppen-Designs aufgrund ethischer und fachlicher Überlegungen nur begrenzt möglich (vgl. Dollinger, 2018, S. 250; Ziegler, 2018).

5. Ausblick auf ein laufendes Forschungsprojekt

Die aktuelle Evaluationspraxis von Organisationen der Sozialen Arbeit wie auch Fragen über deren mögliche Angleichung oder Variation sind in der Sozialen Arbeit im deutschsprachigen Raum noch wenig erforscht. Das Verhältnis von Mechanismen der Angleichung, wie sie der Neoinstitutionalismus beschreibt, und Faktoren, welche organisationsintern die Realisierung von Evaluationsaktivitäten erschweren bzw. begünstigen, ist empirisch weitgehend unbestimmt. Dies gilt auch für die übergreifende Frage, ob Evaluationsaktivitäten primär der Legitimierung dienen und deren Umsetzung und Nutzung von der Ebene der konkreten Leistungserbringung weitgehend entkoppelt sind (Walgenbach & Meyer, 2008), oder ob das Instrumentarium von Evaluationen im Dienste des Wissensaufbaus in Sozialen Organisationen (Baker et al, 2010) steht oder gar zur Schaffung einer Evidenzbasis für in der Praxis selbst entwickelte professionelle Problemlösungen genutzt wird (vgl. Lee & McMillen, 2017, S. 21).

Dieses Forschungsdesiderat adressiert das im Mai 2020 gestartete und vom Schweizerischen Nationalfonds (SNF) geförderte Forschungsprojekt «Auf dem Weg zur 'evidenzbasierten Praxis'? Evaluationen in Organisationen der Sozialen Arbeit». Das Projekt bedient sich hierbei sowohl quantitativer (Online Survey) als auch qualitativer Forschungsmethoden (Fallanalysen), wobei drei Kernfragen im Fokus stehen: 1. Wie sieht die bestehende Evaluationspraxis in Einrichtungen der

Sozialen Arbeit innerhalb der Deutschschweiz aus? 2. Welche organisationsinternen und -externen Faktoren nehmen Einfluss auf die Evaluationspraxis? 3. Wie gestaltet sich die konkrete praktische Umsetzung von Evaluationsaktivitäten in ausgewählten Einrichtungen der Sozialen Arbeit?

Eine Online-Befragung von Leitungspersonen von Einrichtungen der Sozialen Arbeit in der Deutschschweiz bildet die empirische Basis für die Bearbeitung der ersten beiden Kernfragen. Von Interesse sind unterschiedliche Formen von Evaluationsaktivitäten, deren Verschränkung mit professionellen Handlungsvollzügen und die Identifikation verschiedener «Typen der Evaluierung», die für eine bestimmte Evaluationspraxis charakteristisch sind. Mit der Beantwortung der zweiten Kernfrage werden einerseits jene oben diskutierten Einflüsse des Umfeldes auf die bestehende Evaluationspraxis aufgegriffen (siehe Abschnitt 3), welche aus neoinstitutionalistischer Perspektive die treibende Kraft von Angleichungsprozessen darstellen: Isomorphismus durch Zwang (z.B. gesetzliche Vorgaben), normativer Druck (z.B. Stellenwert von Konzepten wie „evidenzbasierte Praxis“) sowie mimetische Prozesse (z.B. Orientierung an Strukturen und Praxis der Evaluation anderer Organisationen im gleichen Arbeitsfeld). Andererseits werden die unter «evaluation capacity» gefassten, organisationsinternen Faktoren (z.B. personelle und finanzielle Ressourcen, Evaluationskultur, Fähigkeiten der Mitarbeitenden) einbezogen (siehe Abschnitt 4) und als potenzielle Erklärung für die Variation der Evaluationspraxis bzw. die Unterscheidung verschiedener «Typen der Evaluierung» eingesetzt. Eine vertiefte Untersuchung der konkreten Evaluationspraxis und insbesondere der Verschränkung bzw. der Entkoppelung von Evaluationsaktivitäten und der professionellen Praxis als Ort der Erzeugung und Nutzung von Wissen findet in einem weiteren Schritt (Kernfrage 3) im Rahmen der Fallanalysen in ausgewählten Einrichtungen der Sozialen Arbeit statt.

Im März 2021 ist der Online-Survey gestartet und wird damit bald erste Analysen zur Evaluationspraxis in Organisationen der Sozialen Arbeit, deren Zusammenspiel mit Bedingungen und Einflussfaktoren und damit Antworten auf hier aufgeworfene Fragen ermöglichen.

Prof. Dr. Edgar Baumgartner, Leiter des Instituts Professionsforschung und -entwicklung an der Hochschule für Soziale Arbeit der Fachhochschule Nordwestschweiz. Themenschwerpunkte: Betriebliche Soziale Arbeit, Evaluationsforschung und Methoden der Wirkungsforschung.

Aline Kaufmann, MSc, Wissenschaftliche Assistentin im Institut Professionsforschung und -entwicklung an der Hochschule für Soziale Arbeit der Fachhochschule Nordwestschweiz. Themenschwerpunkte: Methoden der qualitativen Sozialforschung, Evaluationsforschung, Organisationsentwicklung, Arbeitsmarkt und Gender.

Michael Lambertus, MSc, Wissenschaftlicher Mitarbeiter im Institut Professionsforschung und -entwicklung an der Hochschule für Soziale Arbeit der Fachhochschule Nordwestschweiz. Themenschwerpunkte: Methoden der quantitativen Sozialforschung und Evaluationsforschung.

Literatur

Alaimo, Salvatore P. (2008). Nonprofits and evaluation: Managing expectations from the leader's perspective. In: *New Directions for Evaluation*, (111), 73-92.

Baier, Florian; Keller, Samuel; Koch, Martina & Wigger, Annegret (2013). Die schweizerische Forschungslandschaft der Sozialen Arbeit. In: Riedi, Anna Maria; Zwilling, Michael; Meier

- Kressig, Marcel; Benz Bartoletta, Petra & Aebi Zindel, Doris (Hrsg.). Handbuch Sozialwesen Schweiz. Bern: Haupt, 409–418.
- Baker, Lisa R.; Stephens, Frederick & Hitchcock, Laurel (2010). Social work practitioners and practice evaluation: How are we doing? In: *Journal of Human Behavior in the Social Environment*, 20(8), 963–973.
- Baumgartner, Edgar (2016). Wirkungsmessung in der Sozialen Arbeit in der Schweiz. In: *Soziale Arbeit*, 65 (6/7), 242–248.
- Baumgartner, Edgar & Sommerfeld, Peter (2022, im Erscheinen). Evaluation und evidenzbasierte Praxis. In: Thole, Werner (Hrsg.). *Grundriss Soziale Arbeit. Ein einführendes Handbuch*. Wiesbaden: VS Verlag.
- Becker-Lenz, Roland & Baumgartner, Edgar (2016). Die Profession der Sozialen Arbeit am Beispiel der Schweiz. In: von Dick, Michael; Marotzki, Winfried & Mieg, Harald (Hrsg.). *Handbuch Professionsentwicklung*. Stuttgart: UTB Verlag, 526–535.
- Bernet, Pius & Gmür, Markus (2015). Leistungs- und Erfolgsmessung in Sozialen und Gesundheits-NPO. In: *Verbandsmanagement* (2), 23–33.
- Bestvater, Hanne & Beywl, Wolfgang (2015). Gelingensbedingungen von Selbstevaluation. In: Bolay, Eberhard; Iser, Angelika & Weinhardt, Marc (Hrsg.). *Methodisch Handeln – Beiträge zu Maja Heiners Impulsen zur Professionalisierung der Sozialen Arbeit*. Wiesbaden: Springer VS, 133–145.
- Bieri, Oliver & Studer, Sibylle (2016). Wirkungsmessung bei NPO: Notwendiges Übel oder Teil der Strategie? Wirkungsmessung und Entwicklungspotenzial bei Spenden sammelnden Organisationen in der Schweiz. In: *Soziale Arbeit*, 65(6/7), 267–273.
- Carman, Joanne G. & Fredericks, Kimberly A. (2010). Evaluation Capacity and Nonprofit Organizations. Is the Glass Half-Empty or Half-Full? In: *American Journal of Evaluation*, 31(1), 84–104.
- Cousins, Bradley J.; Goh, Swee C; Elliott, Catherine J & Bourgeois, Isabelle (2014). Framing the capacity to do and use evaluation. In: Cousins, Bradley J; Goh, Swee C; Elliott, Catherine J & Bourgeois, Isabelle (Hrsg.). *Organizational capacity to do and use evaluation. New Directions for Evaluation*, 141, 7–23.
- DiMaggio, Paul J. & Powell, Walter W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. In: *American Sociological Review* (2) 48, 47–60.
- DiMaggio, Paul J. & Powell, Walter W. (1991). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. In: DiMaggio, Paul J. & Powell, Walter W. (Eds.). *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press, 63–82.
- Dollinger, Bernd (2018). Paradigmen sozial- und erziehungswissenschaftlicher Wirkungsforschung: Eine Analyse kausaltheoretischer Annahmen und ihrer Folgen für die Soziale Arbeit. In: *Soziale Passagen*, 20, 245–246.
- Falk, Wiebke (2016). Deinstitutionalisieren durch organisationalen Wandel. Selbstbestimmung und Teilhabe behinderter Menschen als Herausforderung für Veränderungsprozesse in Organisationen. Bad Heilbrunn: Verlag Julius Klinkhardt.

- Froncek, Benjamin; Mazziotta, Agostino; McCann Pamela, J.; Würpel, Ilka & Rohmann, Anette (2018). Evaluation in der psychosozialen Beratungsarbeit: Ein Survey zur Praxis der Evaluation aus Sicht der Beratenden. In: Zeitschrift für Evaluation, 17(2), 241-268.
- Gambrill, Eileen (1999). Evidence-based practice: An alternative to authority-based practice. In: Families in Society, 80(4), 341-350.
- Gervin, Derrick W.; Davis, Sarita K.; Jones, Jenny L.; Counts-Spriggs, Margaret S. E. & Farris, Kiberly D. (2010). Evaluation Development and Use in Social Work Practice. In: Journal of MultiDisciplinary Evaluation, 6(14), 85-101.
- Haubrich, Karin (2008). Evaluation in der Sozialen Arbeit in Deutschland. Entwicklungslinien und Besonderheiten der Evaluationsdebatte am Beispiel der Kinder-, Jugend- und Familienhilfe. In: Widmer, Thomas; Beywl, Wolfgang & Fabian, Carlo (Hrsg.). Evaluation. Ein systematisches Handbuch. Wiesbaden: VS Verlag für Sozialwissenschaften, 441-457.
- Heiner, Maja (2001). Planung und Durchführung von Evaluationen - Anregungen, Empfehlungen, Warnungen. In: Heil, Karolus; Heiner, Maja & Feldmann, Ursula (Hrsg.). Evaluation sozialer Arbeit. Freiburg i. Br.: Lambertus, 35-58.
- Hense, Jan & Widmer, Thomas (2013). Ein Überblick zum internationalen Stand der Forschung über Evaluation. In: Hense, Jan; Rädiker, Stefan; Böttcher, Wolfgang & Widmer, Thomas (Hrsg.). Forschung über Evaluation. Bedingungen, Prozesse, Wirkungen. Münster: Waxmann-Verlag, 251-278.
- Hochuli Freund, Ursula; Gebert, Jakin; John, Anna & Sprenger, Raphaela (2019). «Wissen in der Praxis». Projektbericht. Olten: Institut Professionsforschung und -entwicklung, Hochschule für Soziale Arbeit FHNW.
- Hornbostel, Stefan (2012). Evaluation und Evaluationsforschung. In: Maasen, Sabine; Kaiser, Mario; Reinhart, Martin & Sutter, Barbara (Hrsg.). Handbuch Wissenschaftssoziologie. Wiesbaden: VS-Verlag, 277-287.
- James, Sigrid (2016). «Inside the belly oft he beast» - Möglichkeiten und Grenzen der evidenzbasierten Praxis. In: Borrmann, Stefan; Thiessen, Barbara (Hrsg.). Wirkungen Sozialer Arbeit. Potentiale und Grenzen der Evidenzbasierung für Profession und Disziplin. Opladen, Berlin, Toronto: Verlag Barbara Budrich, 143-160.
- Lee, Bethany R. & McMillen, J. Curtis (2017). Pathways Forward for Embracing Evidence-Based Practice in Group Care Settings. In: Journal of Emotional and Behavioral Disorders, 25(1), 19-27.
- Mayrhofer, Hemma (2014). Krise des Widerspruchsmanagements - Wirkungsorientierung in der Sozialen Arbeit aus neoinstitutionalistischer Theorieperspektive. In: Österreichische Zeitschrift für Soziologie, 39(4), 281-305.
- Merchel, Joachim (2015). Ohne Evaluation kein professionelles Handeln!? Maja Heiners beharrliches Plädoyer für datenbasierte Reflexion in der Sozialen Arbeit. In: Bolay, Eberhard; Isler, Angelika & Weinhardt, Marc (Hrsg.). Methodisch Handeln - Beiträge zu Maja Heiners Impulsen zur Professionalisierung der Sozialen Arbeit. Forschung und Entwicklung in der Erziehungswissenschaft. Wiesbaden: Springer VS, 149-161.
- Meyer, John W. & Rowan, Brian (1977). Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, 83(2), 340-363.

- Meyer, John W. & Rowan, Brian (2009). Institutionalisierte Organisationen. Formale Struktur als Mythos und Zeremonie. In: Koch, Sascha & Schemmann, Michael (Hrsg.). Neo-Institutionalismus in der Erziehungswissenschaft: Grundlegende Texte und empirische Studien. Wiesbaden: VS-Verlag, 28-56.
- Müller, Wolfgang Carl (Hrsg.) (1978). Begleitforschung in der Sozialpädagogik. Analysen und Berichte zur Evaluationsforschung in der Bundesrepublik. Weinheim, Basel: Beltz.
- Naccarella, Lucio; Pirkis, Jane; Kohn, Fay; Morley, Belinda; Burgess, Philip & Blashki, Grant (2007). Building evaluation capacity: Definitional and practical implications from an Australian case study. In: Evaluation and Program Planning, (30), 231-236.
- Norton, Sophie; Milat, Andrew; Edwards, Barry & Giffin, Michael (2016). Narrative review of strategies by organizations for building evaluation capacity. In: Evaluation and Program Planning, (58), 1-19.
- Otto, Hans-Uwe & Ziegler, Holger (2011). Managerialismus. In: Otto, Hans-Uwe & Thiersch, Hans (Hrsg.). Handbuch Soziale Arbeit, 4., völlig. neu bearb. Aufl. München: Ernst Reinhardt Verlag, 901-911.
- Petersén, Anna Charlotta & Olsson, Jan Ingvar (2014). An evaluation paradox in social work? An empirical study of evaluation use in connection with temporary programmes in Swedish social work, In: European Journal of Social Work, 17(2), 175-191.
- Phineo (2013). Wirkungsorientierte Steuerung in Non-Profit-Organisationen. Berlin: PHINEO AG.
- Polutta, Andreas (2014). Wirkungsorientierte Transformation der Jugendhilfe. Ein neuer Modus der Professionalisierung Sozialer Arbeit? Wiesbaden: Springer.
- Polutta, Andreas (2010). Wirkungsorientierung und Profession. Neue Professionalisierung oder Ende der Profession Sozialer Arbeit. In: Soziale Passagen. Journal für Empirie und Theorie Sozialer Arbeit, 2(1), 47-62.
- Preskill, Hallie & Boyle, Shanelle (2008). A multidisciplinary model of evaluation capacity building. In: American Journal of Evaluation, 29(4), 443-459.
- Sackett, David L.; Rosenberg, William M.C.; Gray, J.A.M.; Haynes, R. Brian & Richardson, W. Scott (1996). Evidence based medicine: what it is and what it isn't. In: British Medical Journal, 312, 71-72.
- Schneider, Armin (2016). Konzepte der Wirkungsmessung und -forschung. Zwischen Goldstandard und vergoldeten Attrappen. In: Soziale Arbeit, 65(6/7), 204-211.
- Schulze-Krüdener, Jörgen (2017). Wissen, was in der sozialen Arbeit wirkt! Zur Reichweite empirischer Zugänge. Bremen: Apollon University Press.
- Scott, Richard W. (2001). Institutions and Organizations. Second Edition. Thousand Oaks: Sage Publications.
- Spiegel, Hiltrud von (2008). Evaluation. In: Kreft, Dieter & Mielenz, Ingrid (Hrsg.). Wörterbuch Soziale Arbeit. Aufgaben, Praxisfelder, Begriffe und Methoden der Sozialarbeit und Sozialpädagogik. 6. überarb. und aktualisierte Aufl. Weinheim, München: Beltz, 275-278.
- Stegmann, Michael & Schwab, Jürgen E. (2012). Evaluieren und Forschen für die Soziale Arbeit. Ein Arbeits- und Studienbuch - Reihe Hand- und Arbeitsbücher (H 4). Berlin: Eigenverlag des Deutschen Vereins für öffentliche und private Fürsorge e.V.
- Stockmann, Reinhard (2018). Der Evaluation-Capacity-Development-Ansatz des Centrums für Evaluation (CEval). In: Zeitschrift für Evaluation, 18(2), 363-392.

- Stockmann, Reinhard & Meyer, Wolfgang (2020). Die Institutionalisierung der Evaluation in Europa. In: Zeitschrift für Evaluation, 19(1), 111–135.
- Trenkmann, Jeannette (2017). Soziologischer Neo-Institutionalismus als theoretischer Bezugsrahmen. In: Trenkmann, Jeannette (Hrsg.). Frauen in Führungspositionen der deutschen Privatwirtschaft. Eine institutionensoziologische Analyse organisationaler Gleichstellungspolitik. Wiesbaden: Springer Verlag, 61–107.
- von Spiegel, Hiltrud (2008). Evaluation. In: Kreft, Dieter & Mielenz, Ingrid (Hrsg.). Wörterbuch Soziale Arbeit. Aufgaben, Praxisfelder, Begriffe und Methoden der Sozialarbeit und Sozialpädagogik. 6. überarb. und aktualisierte Aufl. Weinheim, München: Beltz, 275–278.
- von Werthern, Anna (2019). Theoriebasierte Evaluation. Entwicklung und Anwendung eines Verfahrensmodells zur Programmtheoriekonstruktion. Wiesbaden: Springer Fachmedien.
- Walgenbach, Peter & Meyer, Renate (2008). Neoinstitutionalistische Organisationstheorie. Stuttgart: Kohlhammer.
- Widmer, Thomas; Eberli, Daniela; Ackermann, Günter; Balzer, Lars; Bundi, Pirmin; Hirschi, Christian; Kasper, Tanya; Neuenschwander, Peter; Neuhaus, Björn; Rohrbach, Walter; Schlauffer, Caroline & Strebel, Felix (2016). Forschung über Evaluation in der Schweiz: Stand und Aussichten. In: Leges, 27(3), 459–483.
- Ziegler, Holger (2012). Wirkungsforschung – über Allianzen von Evaluation und Managerialismus und die Möglichkeit erklärender Kritik. In: Schimpf Elke & Stehr Johannes (Hrsg.). Kritisches Forschen in der Sozialen Arbeit. Wiesbaden: VS Verlag für Sozialwissenschaften, 93–105.

Sigrid Haunberger & Tobias Jörg

Was wirkt für wen, wie und unter welchen Bedingungen? Wirkungsmodelle, Wirkfaktoren und Wirkungen in der Bewährungshilfe im Blick

1. Ausgangslage

Bis zum heutigen Zeitpunkt wird die Frage nach Wirkungsnachweisen oder «What Works?» in der Sozialen Arbeit kontrovers diskutiert und ist häufig normativ besetzt. Befürworter*innen und Zweifler*innen vereint der Sammelband von Otto et al. (2010), der die Debatte wieder schwingungsbend angestoßen hat. In einem weiteren umfassenden Sammelband von Eppler et al. (2011) kommt zum Ausdruck, wie unterschiedlich Wirkung gedacht und erforscht werden kann. Richtet man den Blick auf wirkungsfeststellende Forschungs- und Evaluationsvorhaben ist erstens auffällig, dass die Soziale Arbeit nach Handlungsfeldern unterschiedlich intensiv und in verschiedenartigen Konnotationen mit der Anforderung konfrontiert wird, die Wirksamkeit ihres Handelns nachzuweisen (Merchel, 2015, S. 128). Zweitens steht sehr häufig die Bestimmung des sog. Outcomes, also der subjektiven Wirkung bei den Adressat*innen oder auch der Impact, die aggregierten Auswirkungen auf die Gesellschaft, im Vordergrund (Macseanaere, 2013, S. 212). Nur vereinzelt findet eine durchaus sinnvolle Erweiterung der Wirkungsforschung statt, indem nicht nur untersucht wird, welche Effekte bei welcher Hilfeart und in welchem pädagogischen Setting erreicht werden, sondern auch, was die hierfür begünstigenden oder hemmenden Faktoren sind (Macseanaere, 2013, S. 214). Die Beschäftigung mit Wirkfaktoren ist bisher bekannt aus der Psychotherapieforschung (Grawe, 2000) und aus der Forschung über Coaching (Wechsler, 2012), stellt aber innerhalb der Sozialen Arbeit und der Bewährungshilfe ein Desiderat dar. Drittens kommt einer Wirkungsforschung, die sich mit Kontextbedingungen und Mechanismen auseinandersetzt und gerade für die Entwicklung professionellen Handelns wichtig wäre, eine bisher unterschätzte Bedeutung zu (Polutta, 2013, S. 195).

Die Debatte über effektive Interventionen für Straftäter*innen hat eine lange Geschichte. Auf der einen Seite scheinen einige Schlüsselprinzipien für effektive Programme aufgrund von meta-analytischen Studien und Literaturübersichten abgesichert und im Strafvollzug gut etabliert zu sein. Auf der anderen Seite wird beanstandet, dass bisher nur sehr wenig Wirkungsforschung über Einzelfallarbeit/Case Management im Allgemeinen und im Speziellen in der Bewährungshilfe vorliegt (Raynor et al., 2014). Die wenigen vorliegenden Studien weisen auf die Bedeutung und speziellen Merkmale der Arbeitsbeziehung zwischen Bewährungshelfer*in und Betreuten hin (Barry, 2000; Durnescu, 2012).

Der hier vertretene Ansatz einer kontextmechanismusbasierten Wirkungsforschung geht davon aus, dass Unterstützungsprozesse den Kern der Praxis Sozialer Arbeit bilden und alle Versuche einer Wirkungsmessung folglich auf der Mikroebene beginnen müssen. Gemäß Haller (2011, S. 238) konstituieren sich diese Unterstützungsprozesse auf der Mikroebene durch a) die entsprechende Lage der Adressat*innen mit ihren Ressourcen und Defiziten, b) den durch die Fachpersonen erbrachten Interventionen und den damit verbundenen Interaktionen zwischen Fachpersonen und Adressat*innen, c) die im Unterstützungsprozess erzielten Wirkungen, d) einem organisatorisch-institutionellen Rahmen, in dem die Unterstützungsprozesse stattfinden (Mesoebene) und e) einem gesellschaftlichen und politischen Rahmen, der Auswirkungen auf die

Lage der Adressat*innen und die institutionellen und fachlichen Bedingungen der Unterstützungsprozesse hat (Makroebene).

Die Behandlung des Themenkomplexes Wirkungsmodelle, Wirkfaktoren und Wirkungen soll hier anhand von zwei konkreten Projekten erfolgen. Zuerst wird ein Forschungsprojekt vorgestellt, welches zusammen mit Praxispartnerorganisationen Wirkungsmodelle für Unterstützungsprozesse entwickelte. Danach wird ein weiteres Projekt vorgestellt, indem eine neue interne Statistik innerhalb des Bewährungsdienstes Thurgau entwickelt wurde, die Wirkungsziele auf Adressat*innenebene verstärkt in den Blick nimmt. Abschließend geht es um die Frage, welche Erkenntnisse sich daraus für eine Wirkungsorientierung in der Bewährungshilfe und eine zukünftige Wirkungsforschung in der Sozialen Arbeit ableiten lassen.

2. Unsere Projekte

2.1 Projekt WIFASO

Das vom Schweizerischen Nationalfonds geförderte anwendungsbezogene Forschungsprojekt 'Wirkfaktoren professionellen Handelns in sozialen Organisationen' (WIFASO) wird im Zeitraum 2018 bis 2021 in Kooperation mit über 20 Deutschschweizer Praxisorganisationen durchgeführt. Die beteiligten Organisationen stammen aus den sozialarbeiterischen Handlungsfeldern: Arbeitsintegration, Bewährungshilfe, Schuldenberatung und Sozialhilfe. Allen Handlungsfeldern ist gemeinsam, dass es um Unterstützungsprozesse zur Förderung der beruflichen, wirtschaftlichen und sozialen Integration der entsprechenden Adressat*innen geht und jeweils die Wirksamkeit dieser auf der Mikroebene im Einzelfall in den Blick genommen wird. Pro Handlungsfeld werden kontextmechanismusbasierte Wirkungsmodelle erstellt, die sich auf unterschiedliche Wissensquellen abstützen. Um den Wissenstand bezüglich Wirkfaktoren und Wirkungen möglichst vollständig abzudecken, wurde für das Projekt ein komplexes multi-methodisches sequentielles Forschungsdesign gewählt (Häder, 2010). Zunächst wurden systematische Literaturübersichten pro Handlungsfeld (vgl. Sundermann et al., 2021; Haunberger et al., 2020) erstellt, aufbauend darauf wurden in einer qualitativen Teilstudie Fachpersonen gebeten, Prozessverläufe in Fällen zu schildern und hinsichtlich erfolgreicher oder weniger erfolgreicher Fallabschlüsse einzuschätzen. Aufbauend hierauf wurde eine quantitative Onlineumfrage durchgeführt, in der Fachpersonen aus den genannten Handlungsfeldern fallbezogen Wirkfaktoren und Wirkungen einschätzen sollten. In einem abschließenden Delphi-Verfahren wurden die entstandenen Wirkungsmodelle nochmals einem Expert*innenblick der am Projekt beteiligten Fachpersonen unterzogen, um die Angemessenheit, Praxistauglichkeit und konkrete Verwendungszwecke zu bestimmen. Weiterführende Informationen sowie aktuelle Publikationen befinden sich auf der Projekthomepage: www.zhaw.ch/wifaso. In diesem Beitrag berichten wir ausschließlich vom Handlungsfeld Bewährungshilfe und geben einen Einblick in das Wirkungsmodell aus dem qualitativen Teilprojekt.

Realist Evaluation

Um die zugrunde liegende Fragestellung zu beantworten, was wann, für wen und unter welchen Kontextbedingungen wirkt, strukturieren wir diesen Beitrag nach der realistischen Evaluation (Pawson & Tilley, 1997), einem vielversprechenden Ansatz für komplexe soziale Interventionen (z.B. Kazi 2003; Kazi et al. 2011; Haunberger & Baumgartner, 2017).

Damit verschiebt sich grundlegend der Schwerpunkt gängiger Fragestellungen in der 'What Works' Debatte bzw. in wirkungsfeststellenden Forschungs- und Evaluationsprojekten. In diesem Beitrag geht es in erster Linie nicht um eine Wirksamkeitsprüfung (im Sinne einer Outcome-Forschung). Es wird nicht gefragt, ob Soziale Arbeit/ Bewährungshilfe wirkt, sondern welche Wirkfaktoren für einen erfolgreichen Unterstützungsprozess unerlässlich sind. Dies schließt selbstverständlich Outcomes nicht aus, verlagert aber den Fokus der Betrachtung auf Bedingungen, die dafür notwendig sind – in der Terminologie der realistischen Evaluation Kontexte und Mechanismen. Kontext-Mechanismus-Outcome-Konfigurationen (KMO) bilden das Herzstück des realistischen Evaluationsansatzes. Sie setzen sich aus drei Schlüsselementen zusammen, die nachfolgend kurz erläutert werden.

Mechanismen rufen in einem bestimmten Kontext bestimmte Ergebnisse hervor. Der Begriff 'generativer Mechanismus' bezieht sich auf die zugrunde liegenden sozialen oder psychologischen Triebkräfte, die die Reaktion der Akteur*innen verursachen.

Kontexte sind notwendig, damit eine Intervention Mechanismen auslöst, die wiederum bestimmte Ergebnisse hervorrufen. Die jeweiligen Kontexte beeinflussen die Argumentation der Akteur*innen. Kontexte können alternative Erklärungen für die beobachteten Ergebnisse bereitstellen. Unterschiedliche Kontexte (z.B. Gesundheitszustand, Familie, Lebensstil, Gemeinschaftsumfeld, Familienzusammensetzung, Subkultur, wirtschaftliche und berufliche Umstände) können eine unterschiedliche Reaktion (Mechanismus) auslösen (Westthorp, 2014, S. 6).

Outcomes umfassen die praktischen Auswirkungen von generativen Mechanismen, die in einem bestimmten Kontext ausgelöst werden (Pawson & Tilley, 1997, S. 412f.).

Für die konkrete Umsetzung einer realistischen Evaluation gibt es keine genauen methodischen Vorgaben. Im Rahmen von systematischen Literaturreviews wurde festgestellt, dass Autor*innen auf ein breites Methodenrepertoire, überwiegend aus dem qualitativen Bereich, zurückgreifen (Haunberger & Baumgartner, 2017).

2.2 Projekt Interne Statistik

Der erste Input zum Projekt für die Implementierung der internen Statistik beim Bewährungsdienst (BWD) Thurgau entstand aufgrund der Auseinandersetzung im Statistikmodul des Masterstudienganges an der Zürcher Hochschule für Angewandte Wissenschaften. In der Vorlesung wurde auf die Qualitätskriterien von empirischer Forschung nach Döring & Bortz (2016, S. 184) hingewiesen sowie die Möglichkeiten und Chancen von empirischen Daten erläutert. Im Zuge dieser theoretischen Auseinandersetzung erfolgte die Analyse des bestehenden internen Controllings/Statistik beim BWD Thurgau, wobei unterschiedliche Defizite ausgemacht werden konnten, welche sich auf die Validität, Reliabilität und Objektivität der bisher erhobenen Daten der durch ein Mandat betreuten Klient*innen bezogen. Die Defizite zeigten sich in der Art der Datenerhebung, der Datenauswertung sowie der Datenpräsentation. Des Weiteren bestand für das interne Controlling keine theoretischen Grundlagen. Es fehlte der Bezug zum empirischen Wissen, die Verortung in einem spezifischen Kontext und die Anschlussfähigkeit zu bestehenden internen Konzeptionen und Theorien. Des Weiteren wurde erkannt, dass durch die regelmäßige und präzise Erfassung der Lebenssituation von Klient*innen Veränderungen im Mandatsverlauf erfasst werden können. Zu den relevanten Merkmalen der Lebenssituation gehören das soziale Umfeld, Familie, Partnerschaft, Schule, Arbeit, Substanzkonsum und Freizeit (Andrews & Bonta, 2010).

Interventionen innerhalb dieser Faktoren können zur Senkung des Rückfallrisikos beitragen. Die Veränderungen innerhalb der erwähnten Faktoren sollen zukünftig erfasst und bezüglich der Wirksamkeit der Unterstützungsprozesse im BWD Thurgau interpretiert werden. Anhand dieser Erkenntnisse kann anschließend das Dienstleistungsangebot des BWD Thurgau angepasst und ausgebaut werden.

In der Projektvereinbarung wurden als Projektziel die folgenden Punkte definiert:

- Die Erarbeitung einer fundierten theoretischen Basis, welche auch mit bestehenden internen Richtlinien und Hilfsmitteln kompatibel ist.
- Die Erstellung einer nominalorientierten Statistik, welche Aussagen zur aktuellen Lebenssituation von Klient*innen des Bewährungsdienstes Thurgau zulässt.
- Die technische Implementierung im bestehenden Dokumentationsprogramm.
- Die Einführung und Schulung der Mitarbeitenden in der Umsetzung und Anwendung der Statistik.

3. Ergebnisse

3.1 Projekt WIFASO

Im Folgenden wird das kontextmechanismusbasierte Wirkungsmodell für das Handlungsfeld Bewährungshilfe vorgestellt, welches im Projekt entwickelt wurde. Grundlage des Wirkungsmodells sind problemzentrierte Interviews, die mit 16 Bewährungshelfer*innen in der Deutschschweiz geführt wurden. Die Bewährungshelfer*innen waren aufgefordert in den Interviews jeweils zwei spontane Fälle, einer mit positivem Fallabschluss (Outcome), einer mit negativem Outcome, zu beschreiben. Damit resultierten insgesamt 32 Fallbeschreibungen, die in das kontextmechanismusbasierte Wirkungsmodell eingeflossen sind. Die zentralen Kontexte, Mechanismen und Outcomes wurden kodiert und inhaltsanalytisch ausgewertet.

Wir nehmen an dieser Stelle mehr auf die Konstruktion von kontextmechanismusbasierten Wirkungsmodellen Bezug und diskutieren die sich darin zeigenden inhaltlichen Ergebnisse nur ansatzweise. Schon an dieser Stelle lässt sich festhalten, dass ein kontextmechanismusbasiertes Wirkungsmodell einer anderen Logik folgt, als es herkömmliche Modellierungen im Sinne von Wirkungsketten tun (Input – Output – Outcome – Impact). Das Wirkungsmodell integriert verschiedene Ebenen miteinander. Zum einen, gemäß realistischer Evaluation, werden Kontexte, Mechanismen und Outcomes erfasst, die zusammengenommen sogenannte KMO-Konfigurationen ergeben. Diese können in Form von Hypothesen im realistischen Evaluationszyklus überprüft, empirisch erhoben und weiterentwickelt werden. Zum anderen wird ebenso auf die Mikro-, Meso- und Makroebene Bezug genommen. Der Schwerpunkt des Wirkungsmodells liegt auf der Mikroebene, die im Weiteren nach Merkmalen der Fachperson und der Klient*innen unterschieden wird, sowie nach Kontexten und Mechanismen. Die erzielten Outcomes befinden sich ausschließlich bei den Klient*innen.

Im Folgenden werfen wir einen Blick auf Kontexte und Mechanismen, die seitens Bewährungshelfer*innen und Klient*innen als hilfreiche Bedingungen für einen erfolgreichen oder nicht erfolgreichen Fallabschluss genannt werden. Ebenso nehmen wir Bezug auf die am häufigsten genannten positiven Outcomes (vgl. Abb. 1).

Im Kontext der Bewährungshelfer*innen sind die Vernetzung und Kooperation mit Externen sowie die Identifikation mit der Arbeit und der Organisation Bedingungen, die am häufigsten als

unterstützend genannt werden. Auch die interne Zusammenarbeit, der Austausch im Team, die Klient*innenorientierung und die Flexibilität sind hilfreich für einen erfolgreichen Fallabschluss.

Im Kontext der Klient*innen werden folgende Bedingungen am häufigsten als unterstützend genannt: finanzielle Stabilität, Arbeit / einen Arbeitsplatz haben (als Garantie einer Tagesstruktur), eine stabile Wohnsituation, ein motivierendes soziales Umfeld / Familie als Motivationsquelle.

Bei den unterstützenden Mechanismen auf Seiten der Bewährungshelfer*innen werden der Aufbau der Arbeitsbeziehung und des Vertrauens, Transparenz, Offenheit, Erkennen von Risikofaktoren, Informieren und Aufzeigen von Perspektiven bzw. Konsequenzen am häufigsten als unterstützend genannt, ebenso wie das Setzen von Zielen und die Förderung der Eigenverantwortung.

Am häufigsten genannte unterstützende Mechanismen auf Seiten der Klient*innen sind Motivation bzw. Bereitschaft zur Teilnahme und zum Mitmachen, Zuverlässigkeit, Wunsch nach Veränderung, gefolgt von der Teilnahme an der Therapie, Kooperation und Vertrauen in die Therapeut*innen.

Die drei am häufigsten genannten positiven Ergebnisse sind die erfolgreiche soziale Reintegration (drogen-/alkoholfreies Leben, abgeschlossene Ausbildung, Schuldensanierung und finanzielle Absicherung, Fortführung bzw. Abschluss der Therapie, stabile Wohnsituation, Verbesserung des Gesundheitszustandes), die erfolgreiche berufliche Reintegration (Arbeit finden/halten) und die Problemeinsicht und das Erkennen von Risikofaktoren bei sich selbst.

Hypothesen in Form von KMO-Konfigurationen könnten wie folgt aussehen und beanspruchen keine Kausalität. Es wird ein positiver Fallabschluss und ein negativer Fallabschluss dargestellt.

Fall 1 positiver Fallabschluss: Der Klient geht einer Beschäftigung nach (K), der Klient weist einen ehemaligen Suchtmittelmissbrauch auf (K), der Bewährungshelfer zeigt Perspektiven auf (M), begegnet dem Klienten mit Transparenz und Offenheit (M), der Klient kooperiert im Unterstützungsprozess (M) und sieht einen Nutzen für sich (M), reflektiert den eigenen Suchtmittelmissbrauch (M) und erkennt Verhaltensmuster (M). Hieraus resultiert Problemeinsicht/Risikofaktoren bei sich erkennen (O).

Fall 2 negativer Fallabschluss: Die Zuständigkeit für den Klienten ist ungeklärt (K), die durchgehende Betreuung ist unterbrochen (K), der Klient weist einen bestehenden Suchtmittelmissbrauch auf (K), der Aufbau einer Arbeitsbeziehung ist erschwert (M), der Klient lügt und manipuliert (M). Der Suchtmittelmissbrauch bleibt bestehen (O).

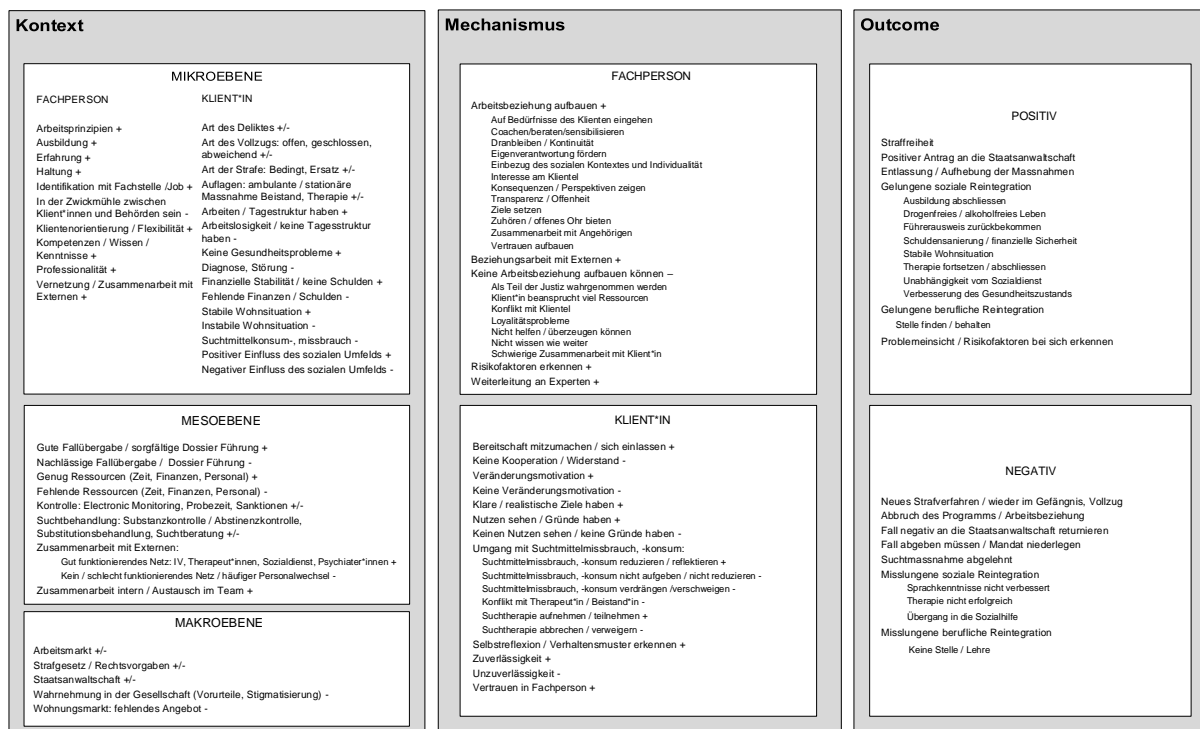


Abbildung 1: Wirkungsmodell Bewährungshilfe aus dem qualitativen Teilprojekt (erstellt von Dr. Zuzanna Kita)

Neben der Sichtbarmachung von zentralen Wirkfaktoren ist für die fachliche Steuerung von Unterstützungsprozessen in der Bewährungshilfe von entscheidender Bedeutung, ob für die Formulierung von Zielen auf Ebene der Adressat*innen bei den Leistungszielen (Output) oder Wirkungszielen (Outcome) angesetzt wird. Sobald eine Messung in Form eines Controllings oder einer Evaluation zur Zielerreichung angestrebt wird, unterscheiden sich der Komplexitäts- und Anforderungsgrad erheblich. Während sich Outputziele (typischerweise sind das Anzahl der Konsultationen, Gesprächsdauer) relativ einfach erfassen lassen, sind an die Formulierung von fachlich begründeten Wirkungszielen und deren Messung weit höhere Ansprüche zu stellen (vgl. Liebig, 2013). Wie eine Operationalisierung von fachlich begründeten Wirkungszielen in der Bewährungshilfe aussehen könnte, wird im Weiteren vorgestellt.

3.2 Projekt Interne Statistik

Für die Statistik des BWD Thurgau wurde als Basis das Risk-Need-Responsivity Modell von Andrews & Bonta (2010) gewählt. Die Entscheidung zugunsten dieser theoretischen Konzeption erfolgte aufgrund dessen, dass das RNR-Modell bereits in der Praxis des BWD Thurgau verankert ist und in überkantonalen Konzepten zur Anwendung kommt. Zur weiteren Differenzierung, insbesondere dem Festlegen der zu erfassenden Hauptkategorien und Items, wurde auf das Prognoseinstrument Level of Service Inventory-Revised (LSI-R) von Andrews & Bonta (2014) zurückgegriffen, mit welchem die kriminogenen Rückfallrisiken in einem «Risk-Need-Assessment» erfasst werden, um anschließend die Rückfallrisiken, die Bedarfsplanung sowie die inhaltliche Ausrichtung der Interventionsplanung zu strukturieren (Dahle et.al., 2012). Als die häufigsten empirisch belegten Prädiktoren gelten hierbei die aus dem RNR-Modell bekannten Risikofaktoren von «antisozialer Einstellungen, antisozialer Sozialkontakte, Merkmale einer antisozialen

Persönlichkeit sowie hieraus erwachsende kumulierende Schwierigkeiten in der Familie, in der Schule, der Arbeit, in der Freizeit oder in der Partnerschaft» (Dahl et al., 2012, S. 7). Die erwähnten Risikomerkmale lassen sich konzeptuell als Faktoren verstehen, die in einer konkreten Handlungssituation die Wahrscheinlichkeit bei Täter*innen für kriminelles Verhalten erhöhen. Zutreffende Risikomerkmale können somit einen deliktfördernden Einfluss auf die Abwägung von Kosten und Nutzen für erneute Delinquenz haben. Im LSI-R werden die erwähnten Risikofaktoren des RNR-Modells in 54 Merkmale aufgeschlüsselt und anschließend in den folgenden zehn Hauptkategorien gefasst:

- 1) Kriminelle Vorgeschichte
- 2) Ausbildung/Erwerbstätigkeit
- 3) Finanzielle Situation
- 4) Familie/Partnerschaft
- 5) Wohnsituation
- 6) Freizeitgestaltung
- 7) Freundschaften/Bekanntschaften
- 8) Alkohol-/Drogenproblematik
- 9) Emotionale/Persönliche Beeinträchtigung
- 10) Einstellung/Orientierung/Werthaltung

Die Auswahl der für den BWD Thurgau relevanten Hauptkategorien ergab sich aus der Unterscheidung von personen- und umweltbezogenem Veränderungsbedarf. Da Themen, welche dem personenbezogenen Veränderungsbedarf zuzuordnen sind, durch psychiatrisch-psychologische Interventionen bearbeitet werden, wurden diese nicht spezifisch in der internen Statistik berücksichtigt. Die Auswahl konzentrierte sich auf die Kategorien, die einer fachlich begründeten Bearbeitung durch die Bewährungshilfe zugänglich sind. Von den zehn Überkategorien des LSI-R wurde somit die Ausbildung/Erwerbstätigkeit, Finanzielle Situation, Wohnsituation sowie die Alkohol-/Drogenproblematik verwendet. Als zusätzliche Kategorie wurde von Seiten der Projektleitung die Straffälligkeit und der Behandlungs-/Betreuungsbedarf miteingeführt. Da die erwähnten Überkategorien häufig einen vielschichtigen Punkt erfassen, wurden diese weiter in den Hauptkategorien aufgeschlüsselt. Der Erhebungsbogen setzt sich zusammen aus acht Hauptkategorien, welche die Lebenssituation der Klient*innen des BWD Thurgau erfassen und auch im Wirkungsmodell in Abb. 1 sichtbar werden. Dies geschieht über die Erwerbssituation (ERW1), der Art des Einkommens (EIN2), dem höchsten Bildungsabschluss (BIL3), der Wohnsituation (WOH4), der finanziellen Situation (FIN5), dem Substanzkonsum (SUB6), der Straffälligkeit (STR7) sowie dem Behandlungs- und Beratungsbedürfnis (BEH8). Die einzelnen Überkategorien bestehen aus sechs bis 13 Items. Im Folgenden wurde zur Veranschaulichung die Hauptkategorie der finanziellen Situation (FIN5) mit den zwölf einzelnen Items abgebildet (vgl. Abb. 2).

Code	Finanzielle Situation (FINS)
1	Keine Schulden
2	Schuldenhöhe < CHF 30'000.00
3	Schuldenhöhe < CHF 70'000.00
4	Schuldenhöhe > CHF 100'000.00
5	Betreibungen vorhanden
6	Schuldensanierung vorhanden
7	Lohnpfändung vorhanden
8	Ratenzahlungen vorhanden
9	Lohnverwaltung vorhanden
-1	Noch unbekannt
-2	Nicht feststellbar
-3	Alternative:

Abbildung 2: Hauptkategorie der finanziellen Situation, Auszug aus der Kodierungstabelle

Die Erfassung innerhalb der Hauptkategorie erfolgt durch die fallführenden Sozialarbeiter*innen bei allen Klient*innen, welche in einem aktuellen strafrechtlichen Mandat durch den BWD Thurgau betreut werden. Als Grundvoraussetzung für die Erfassung gilt eine minimale Betreuungsdauer von einem Jahr. Dies aufgrund dessen, dass Interventionen und dadurch ausgelöste Veränderungen sich häufig erst nach längerer Zeit manifestieren. Durch den Ausschluss von Klient*innen mit einer unterjährigen Betreuungsdauer können Messverzerrungen vermieden oder zumindest reduziert werden. Die Erhebung der Daten erfolgt das erste Mal nach der Falleröffnung. Hierbei sollte das Erstgespräch mit den Klient*innen abgewartet werden, damit aktuelle Angaben berücksichtigt werden können. Sollten Hauptkategorien aufgrund der kurzen Betreuungsdauer noch nicht beantwortet werden können, so kann dies mit dem Item «noch unbekannt» angegeben werden. Das Ziel ist es, dass im weiteren Betreuungsverlauf die noch unbekannt Informationen ergänzt werden. Die nächste Messung erfolgt sechs Monate nach der Falleröffnung. Anschließend wird, ausgehend vom Betreuungsbeginn, alle zwölf Monate eine Erhebung durchgeführt. Die Daten werden auf einem Klientenblatt gesammelt und anschließend durch die Sekretariatsmitarbeitenden des Vollzugs- und Bewährungsdienst Thurgau im internen Daten- und Dokumentationsprogramm eingegeben. Die Auswertung der erhobenen Daten erfolgt in halbjährlichem Rhythmus. Damit liegen auf Fallebene Längsschnittdaten zur Lebenssituation von Klient*innen vor, die Auskunft und Hinweise über die Wirksamkeit der Arbeit der Bewährungshilfe geben können.

4. Schlussfolgerungen und Ausblick

4.1 Wirkungsorientierung in der Praxis der Bewährungshilfe

Primär soll mit der Einführung und Umsetzung der neuen Statistik die Lebenssituation der Klient*innen des Bewährungsdienstes Thurgau erfasst werden. Die Datenerfassung wurde nicht im Sinne einer Kontrollfunktion für Vorgesetztenstellen konzipiert, sondern dient vorrangig dazu, auf Abteilungsebene das Dienstleistungsangebot auszuwerten und verbessern zu können. Die gesammelten Daten dienen zur Orientierung und sollen Veränderungen und Tendenzen in der Lebenssituation der Klient*innen im Betreuungsverlauf sichtbar machen. In Kombination mit dem WIFASO-Wirkungsmodell für die Bewährungshilfe ergibt dies eine argumentative Grundlage gegenüber Vorgesetztenstellen. Benötigte Ressourcen und Unterstützungsangebote können somit in Form eines Wirkungsmodells, Daten und theoretischen Inhalten begründet werden. Des

Weiteren sollen die Inhalte des WIFASO-Wirkungsmodells zukünftig in die Definition der Hauptkategorien und der Items einfließen. Für die Zukunft wäre es wünschenswert, wenn Zusammenhänge von Hauptkategorien (z.B. Finanzielle Situation und erneute Straffälligkeit) empirisch nachgewiesen werden könnten. Ebenso erstrebenswert wäre es, wenn Zusammenhänge von Wirkfaktoren und Wirkungen verstärkter in den Blick genommen werden könnten. In einem nächsten Schritt ist die Erfassung der Zufriedenheit der Klient*innen bezüglich dem Dienstleistungsangebot des BWD Thurgau angedacht. Hierbei soll jährlich mittels Fragebogen eine Rückmeldung zum Betreuungs- und Dienstleistungsangebot eingeholt werden.

Das in diesem Beitrag vorgestellte kontextmechanismusbasierte Wirkungsmodell weist zahlreiche praktische Verwendungsmöglichkeiten auf. Es kann schon jetzt Bewährungshilfen dienen, wenn Unterstützungsprozesse wirkungsorientiert geplant, dokumentiert oder dargestellt werden sollen. Damit können die kritischen Elemente für den Erfolg und die Indikatoren im Kontext von Qualitätssicherung und Wirkungsorientierung identifiziert werden (Haubrich, 2010). Wirkungsmodelle scheinen insbesondere attraktiv zu sein fürs Berichtswesen (Falldokumentation) sowie Qualitätskonzepte und -entwicklung (Liebig, 2013).

4.2 Kontextmechanismusbasierte Wirkungsforschung in der Sozialen Arbeit

Mit dieser qualitativen Teilstudie, der ein realistischer Evaluationsrahmen zugrunde liegt, ist es uns ansatzweise gelungen, die „Black-Box“ von Unterstützungsprozessen in der Bewährungshilfe zu öffnen, und wir können die Frage beantworten: Wie funktioniert das?

Dennoch stellt die realistische Evaluation kein Allheilmittel dar, wesentliche Punkte bleiben ungeklärt und können weitere Forschungsvorhaben stimulieren. Dies betrifft vor allem die Unterscheidung zwischen Kontexten und Mechanismen, die Identifikation von entscheidenden Mechanismen sowie die empirische Überprüfung von KMO-Konfigurationen (vgl. Haunberger, 2015).

Zusammenfassend lässt sich festhalten, dass in den explorativen qualitativen Interviews mit den Bewährungshelfer*innen zahlreiche entscheidende Wirkfaktoren (Kontexte und Mechanismen) sowie Wirkungen (Outcomes) auf Fallebene identifiziert werden konnten. In einem weiteren Schritt können hieraus Hypothesen entwickelt werden, die Kontexte, Mechanismen und Outcomes miteinander verbinden. Wir verstehen dies als Beitrag zu einer Wirkungsforschung, die empirisch fundiertes Erklärungs- und Reflexionswissen auf der Grundlage von Kontext-Mechanismus-Outcome Konfigurationen bereitstellt und damit ermöglicht, empirisch belastbare Theorien mittlerer Reichweite über Wirkungszusammenhänge aufzustellen, mit denen professionelles Entscheidungshandeln (in der Bewährungshilfe) wirkungsorientiert fundiert werden kann (vgl. Otto et al. 2010, S. 20).

4.3 Ausblick

Wirkungsmodelle können also einen Teilaspekt einer wirkungsorientierten Steuerung darstellen. Von einer Orientierung an (fachlich begründeten) Wirkungen profitieren politische Akteur*innen, das Management der leistungserbringenden Organisation sowie die Fachpersonen der Sozialen Arbeit und damit nicht zuletzt die Adressat*innen (Liebig 2013).

Zukünftige Wirkungsevaluationen, die Plausibilitäten aufzeigen wollen, werden im besten Fall von einem Wirkungsmodell begleitet, das Kontexte, Mechanismen und Outcomes bei den

Adressat*innen (theoriegestützt) benennt (Merchel 2015, S. 139-140). Nicht zuletzt stellen seriös erstellte Wirkungsanalysen und -modelle für eine lernende Praxis eine große Chance dar, ihre eigene Fachlichkeit zu reflektieren und bei Bedarf weiterzuentwickeln.

Prof. Dr. Sigrid Haunberger, Dipl. Sozialarbeiterin (FH), Dipl. Soziologin, Dozentin und Projektleiterin an der ZHAW Zürcher Hochschule für Angewandte Wissenschaften, Departement Soziale Arbeit, Institut für Sozialmanagement, Forschungsschwerpunkte und -interessen: Evaluationen und Evaluationsforschung, Wirkungsforschung, quantitative Forschungsmethoden, Soziale Arbeit im Justizvollzug.

Tobias Jörg, Sozialarbeiter (FH), ist seit 2016 beruflich tätig im justiziellen Zwangskontext der Sozialen Arbeit. Berufsbegleitend seit 2018 im MSc in Sozialer Arbeit an der ZHAW Zürcher Hochschule für Angewandte Wissenschaften.

Literatur

- Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed). Lexis Nexis/Anderson Pub.
- Andrews, D. A., & Bonta, J. (2014). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, 16(1), 39–55. <https://doi.org/10.1037/a0018362>
- Barry M. (2000) The mentor/monitor debate in criminal justice:'what works' for offenders. *British Journal of Social Work* 30: 575–595.
- Dahle, K.-P., Harwardt, F., & Schneider,Njepel. (2012). *LSI-R Inventar zur Einschätzung des Rückfallrisikos und des Betreuungs- und Behandlungsbedarfs von Straftätern*. Hogrefe Verlag GmbH.
- Durnescu I. (2012) What matters most in probation supervision: Staff characteristics, staff skills or programme? *Criminology & Criminal Justice* 12: 193–216.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. vollständig überarbeitete, aktualisierte und erweiterte Auflage). Springer.
- Eppler, N., Miethe, I., & Schneider, A. (2011). *Qualitative und quantitative Wirkungsforschung: Ansätze, Beispiele, Perspektiven*. Babara Budrich.
- Grawe, K. (2000). *Psychologische Therapie*. Hogrefe.
- Häder, M. (2010). *Empirische Sozialforschung*. Springer.
- Haller, D. (2011). Wirkungsforschung zur Entwicklung der Professionalität, Identität und Legitimation Sozialer Arbeit. In Eppler, N., Miethe, I., Schneider, A. (Hrsg.), *Qualitative und quantitative Wirkungsforschung: Ansätze, Beispiele, Perspektiven* (S. 235–254). Babara Budrich.
- Haunberger, S., Wyder, A., Gisler, F., & Kita, Z. (2020). Was wirkt in der gesetzlichen Sozialhilfe? Ergebnisse einer systematischen Literaturübersicht. *Blätter der Wohlfahrtspflege*, 6(167), 200–205.
- Haunberger, S. (2015). Realistic Evaluation als Evaluationsrahmen in der Klinischen Sozialarbeit. In G. Hahn & M. Hüttemann (Hrsg.), *Evaluation psychosozialer Interventionen. Jahrbuch Klinische Sozialarbeit VII* (S. 47–58). Bonn: Psychiatrie Verlag.

- Haunberger, S., Baumgartner, E. (2017). Wirkungsevaluationen in der Sozialen Arbeit mittels Realistic Evaluation: empirische Anwendungen und methodische Herausforderungen: eine systematische Literaturreview. *Zeitschrift für Evaluation*, 16(1), 121-145.
- Kazi M. (2003) Realist evaluation for practice. *British Journal of Social Work* 33: 803-818.
- Kazi M., Pagkos B., & Milch HA. (2011) Realist evaluation in wraparound: A new approach in social work evidence-based practice. *Research on Social Work Practice* 21: 57-64.
- Liebig, R. (2013). Wirkungsorientierung in der Kinder- und Jugendarbeit. Anmerkungen zu einer ungeliebten Debatte. In: Land Steiermark – A6 Bildung und Gesellschaft; FA Gesellschaft und Diversität – Referat Jugend (Hrsg.): jugendarbeit: wirkt (S. 67-81). Graz: Verlag für Jugendarbeit und Jugendpolitik.
- Macsenaere, M. (2013). Wirkungsforschung in den Hilfen zur Erziehung. In Grasshoff, Gunther (Eds.), *Adressaten, Nutzer, Agency*, (S. 211-225). Springer.
- Merchel, J. (2015). *Evaluation in der Sozialen Arbeit*. 3 Auflage. UTB.
- Otto, H.-U. (2010). Zum Diskurs um evidenzbasierte Soziale Arbeit. In Otto, H.-U., Polutta, A., & Ziegler, H. (Hrsg.), *What works–welches Wissen braucht die Soziale Arbeit* (S. 7-28). Babara Budrich.
- Pawson R & Tilley N. (1997) *Realistic evaluation*: Sage.
- Polutta, A. (2013). Wirkungsorientierung in der Jugendhilfe. In Grasshoff, G. (Hrsg.), *Adressaten, Nutzer, Agency* (S. 195-210). Springer.
- Raynor P., Ugwudike P., & Vanstone M. (2014) The impact of skills in probation work: A reconviction study. *Criminology & Criminal Justice* 14: 235-249.
- Sundermann, Larissa et al. (submitted). How Do Supported Employment Programs Work? Answers from a Systematic Literature Review. *International Journal for Educational and Vocational Guidance*.
- Wechsler, T. (2012). Das coachingtool-spezifische Wirkfaktorenmodell. *Organisationsberatung, Supervision, Coaching*, 19(4), 405-423.
- Westhorp G. (2014) *Realist impact evaluation: an introduction*. London: Overseas Development Institute: 1-12.

Christoph Liel, Doreen Herler & Edgar Schulz

Selbstevaluation als Mittel zur Wirkungsforschung in der Sozialen Arbeit¹

Eine Pilotevaluation des kanadischen Caring Dads Programms in Deutschland

Wirkung und Wirkungsforschung sind in der Sozialen Arbeit schillernde Begriffe. Der Begriff der Wirkung wird im Duden definiert als „durch eine verursachende Kraft bewirkte Veränderung, Beeinflussung, bewirktes Ergebnis“ (Dudenredaktion o.J.). Hinsichtlich der Frage, inwieweit Wirkungen überhaupt gemessen werden können, und der methodischen Ausrichtung einer Wirkungsforschung in der Sozialen Arbeit gab es viele Kontroversen (Sommerfeld; Hüttemann 2007, Borrmann; Thiessen 2016). Ungeachtet dessen scheint sich auch ein Paradigma der Evidenzbasierung in der Wissenschaft der Sozialen Arbeit durchzusetzen (Engelke et al. 2016). Allerdings bleibt es eine Herausforderung, Ursache-Wirkungs-Zusammenhänge in der Praxis nachzuweisen, da die Maßnahmen der Sozialen Arbeit oft nur einen Aspekt komplexer (bio-)psychosozialer Lebenszusammenhänge ihrer Klientel darstellen. Etablierte Verfahren der Wirkungsforschung, um „störende“ weitere Einflüsse zu kontrollieren, sind praktisch oft nicht umsetzbar oder sehr teuer. Der vorliegende Beitrag plädiert für ein komplexes Verständnis von Wirkungen, indem weitere methodische Möglichkeiten genutzt werden, um nichtkontrollierte Befunde zu validieren und sie kontextuell beziehungsweise international vergleichend zu bewerten. Ziel ist es, in Arbeitsfeldern mit wenig Wirkungswissen zu einer Datenbasis zu gelangen, die den Einstieg in Forschungszyklen und die Kontextualisierung zukünftiger Interventionsforschung ermöglicht. Methodisch stellen quantitative Designs wie im folgenden Praxisbeispiel allerdings nur einen Aspekt der Erforschung von Wirkungen dar.

Der vorliegende Beitrag nutzt ein Praxisbeispiel mit hoher Relevanz für eine Wirkungsforschung der Sozialen Arbeit, nämlich die Arbeit mit Vätern zur Prävention von Kindesmisshandlung. Diese Maßnahmen stehen wie alle Maßnahmen im Kinderschutz und der Gewaltprävention unter einem höheren Erfolgsdruck und der Pflicht zur Legitimierung mittels eines Wirkungsbeleges, verglichen mit Maßnahmen in anderen Arbeitsfeldern. Bisher gibt es kaum zielgerichtete, geschweige denn evaluierte Angebote für Väter im Kontext der Kinder- und Jugendhilfe, weshalb forschungsmethodische Herausforderungen hier besonders offensichtlich werden.

Vor diesem Hintergrund wurde im Rahmen des hier vorgestellten Vorhabens ein kanadisches Präventionsprogramm von Kindesmisshandlung in Deutschland implementiert, noch bevor dieser Ansatz in seiner kulturellen Übertragbarkeit evaluiert worden war. Potentiell unterschiedliche Traditionen, Denkweisen und Sprachgebräuche könnten in beiden Kulturkreisen schließlich zu abweichenden Ergebnissen führen, da sie methodische Grundlagen für die Soziale Arbeit bilden.

Einen Wirkungsbeleg zu erbringen, ist in der Praxis vielfach schwierig, weil methodische Goldstandards wie die randomisierte Zuteilung von Kontroll- und Interventionsgruppen beziehungsweise überhaupt die Bildung von echten Kontrollgruppen aufgrund praktischer

¹ Dieser Beitrag wurde wissenschaftlich begutachtet und erstveröffentlicht in der Zeitschrift *Soziale Arbeit* 12/2021 des Deutschen Zentralinstituts für soziale Fragen (DZI).

Voraussetzungen und ethischer Einwände unmöglich sind. Hinzu kommt, dass die eigentlich interessierten Zielkriterien, wie zum Beispiel die dokumentierte Gewaltrate oder die beobachtete Eltern-Kind-Interaktion, entweder nicht oder nur mit hohem Ressourceneinsatz zugänglich sind. Es ist also nicht verwunderlich, dass die deutsche Kinder- und Jugendhilfe insgesamt noch auf eine geringe Wirkungsforschungstradition zurückblickt, von Ausnahmen abgesehen (zum Beispiel Suess et al. 2010). Hinzu kommt, dass es in der deutschen Sozialarbeitswissenschaft viele konzeptionelle Einwände gegen diese Form der Wirkungsorientierung gibt.

Es ist somit geboten, im Rahmen einer Wirkungsforschung der Sozialen Arbeit Präventionsmaßnahmen zum Kinderschutz und der Stärkung elterlicher Erziehungsfähigkeit zu fokussieren, insbesondere wenn es sich um die gering untersuchte Zielgruppe der Väter handelt.

Wissenschaftlich begleitete Selbstevaluation

Das hier vorgestellte Vorhaben² stand vor der Herausforderung, ein international noch nicht hinreichend wirkungsbelegtes und in seiner Anwendbarkeit im deutschsprachigen Kulturraum ungeprüftes Präventionsangebot in der Kinder- und Jugendhilfe zu implementieren und bei sehr begrenzten Mitteln Wirkungen zu untersuchen. Aus diesem Grund wurde als Evaluationsmethode die Mischform einer Selbstevaluation (König 2007) mit wissenschaftlicher Begleitung gewählt. Der Vorteil der wissenschaftlich begleiteten Selbstevaluation (Liel; Kindler 2009) bestand darin, dass sie schrittweise und zum Teil retrospektiv durchführbar war, nachdem seit der Erprobungsphase in Deutschland (Modellprojekt im Jahr 2008) die weitere Etablierung des Präventionsansatzes von den Erfahrungen und Ressourcen in der Praxis abhing. Gleichzeitig konnte die Praxisentwicklung durch das Einbringen der Forschungsperspektive und des forschungsmethodischen Knowhows insoweit unterstützt werden, dass einem vorläufigen Wirkungsbeleg inklusive vorläufiger Überprüfung der transkulturellen Übertragbarkeit des Ansatzes genüge getan wurde (Liel et al. 2021). Da die Längsschnittstudie im Prä-/Post-Design den methodischen Goldstandard der Wirkungsforschung nicht erfüllen konnte, kamen mit der klinischen Signifikanz und dem Reliable-Change-Index (RCI, Jacobson; Truax 1991) methodische Konzepte aus der Gesundheitsforschung zum Einsatz, welche der Erhärtung von Befunden in nichtkontrollierten Designs dienen. Diese Verfahren sind in der deutschen Wirkungsforschung zur Sozialen Arbeit unüblich.

Prävention von Kindesmisshandlung mit dem Caring Dads Programm

Das kanadische Caring Dads Programm zielt auf Psychoedukation und Wissensvermittlung zu altersangemessenem Erziehungsverhalten und der Entwicklung positiver Beziehungen zum Kind beziehungsweise der Mutter des Kindes ab. Entwickelt im Umfeld der Täterarbeit bei häuslicher Gewalt setzt es an den Überschneidungen von Kindesmisshandlung und Partnergewalt an (Liel 2018) und adressiert dysfunktionales Erziehungsverhalten. Dies scheint erfolgversprechend, da

² Die Pilotevaluation wurde im Rahmen eines Modellprojektes durch den *Landschaftsverband Rheinland* im Jahr 2008 und mit einer Expertise aus Forschungsmitteln des *Nationalen Zentrums Frühe Hilfen* (NZFH) durch das *Bundesministerium für Familie, Senioren, Frauen und Jugend* im Jahr 2016 gefördert. Das *NZFH* ist eine Kooperation der *Bundeszentrale für gesundheitliche Aufklärung* und dem *Deutschen Jugendinstitut e.V.*

sich in der deutschen Täterarbeit gezeigt hat, dass Konflikte zwischen den Eltern über die Kindererziehung zwar die Wahrscheinlichkeit erhöhen, dass in Partnerschaften gewalttätige Väter die Maßnahmen abschließen, Väter mit einem sehr hohen Risiko für Kindesmisshandlung sie aber überzufällig häufig abbrechen (Liel 2017). Ein zielgerichtetes Angebot für diese besonders gefährdete Gruppe von Vätern bei spezifischer Adressierung ihrer Motivation einer verantwortungsbewussten Vaterschaft erscheint somit sinnvoll. Anders als klassische Ansätze der Täterarbeit ist Caring Dads weniger konfrontativ und vielmehr positiv unterstützend ausgerichtet. Die psychoedukativen Elemente zur altersgerechten Entwicklung und Erziehung von Kindern haben Eingang in andere Angebote gefunden und sind insbesondere auch für den selektiven Präventionsbereich³ interessant. Caring Dads ist international nach wie vor das einzige in Ansätzen evaluierte Präventionsangebot für gewalttätige oder gewaltgefährdete Väter (Zimmermann im Druck). Praktische Erfahrungen mit dem Programm liegen aus Kanada, Großbritannien, Norwegen, Australien und Deutschland vor, wobei die hier berichtete deutsche Studie von Liel et al. (2021) erst die dritte nach den kanadischen und britischen Evaluationen von Scott; Lishak (2012) sowie McConnell et al. (2017) ist.

Nach kanadischem Vorbild unterstützt Caring Dads Väter darin, ihr Erziehungsverhalten und die Beziehung zu ihren Kindern dauerhaft zu verbessern. Die teilnehmenden Väter sollen lernen, Konflikte und Krisen in der Familie gewaltfrei zu lösen, elterlich-väterliche Kompetenzen auszubilden und umzusetzen, ihre väterliche Fürsorge zu verbessern, eine gesunde Entwicklung der Kinder zu unterstützen und eine positive und unterstützende Haltung zur Mutter des Kindes zu entwickeln.

Das soziale Gruppentraining umfasst 17 Termine à zwei Zeitstunden/Woche und verfolgt vier Ziele:

1. Väter haben sich mit ihrer Vaterrolle auseinandergesetzt.
2. Väter haben verstanden, was ihre Kinder brauchen und was ihnen schadet.
3. Väter haben sich mit ihrer eigenen Gewalttätigkeit auseinandergesetzt und Verantwortung übernommen.
4. Väter haben gelernt, die Beziehung zu ihren Kindern zu verbessern und zu stärken.

Folgende Schritte sollen zur Erreichung der Ziele bewältigt werden:

1. *Empfinden*: Väter setzen sich mit ihrer Geschichte auseinander: Wie ging es mir mit meinem Vater?
2. *Mitfühlen*: Väter lernen, dass ihre Kinder eigene Bedürfnisse haben.
3. *Erkennen*: Väter stellen sich ihren Fehlern. Sie erkennen, was sie falsch gemacht haben und nehmen die Herausforderung an, ihr Verhalten zu ändern.

³ Gemäß der Zielgruppenorientierung wird Prävention für die Bevölkerung (universell), für Risikogruppen (selektiv) und für Problemträger (indiziert) unterschieden.

4. *Neu handeln*: Väter beginnen aktiv, das Vertrauen ihrer Kinder wieder aufzubauen, trainieren neues Handeln mit dem Ziel einer dauerhaften Verhaltensänderung.

Voraussetzungen für eine Teilnahme sind eine minimale Einsicht des Vaters in eigenes schädigendes Verhalten gegenüber dem Kind und der regelmäßige Vater-Kind-Kontakt beziehungsweise -Umgang (keine Teilnahme bei hochkonflikthaften Sorgerechts- und Umgangsstreitigkeiten). Es wird ein Vertrag/eine Vereinbarung abgeschlossen, welche Informationsweitergabe und Schweigepflicht, Kommunikation und Gruppenregeln festlegt. Zudem gibt es Kontakt der durchführenden Stelle zu den Müttern des Kindes und einen Abschlussbericht sowohl für den Vater als auch für die zuweisende Institution.

Ausgangspunkt für die Einführung von Caring Dads in Deutschland war die Initiative von Fachkräften der Diakonie Düsseldorf, mit einem neuen Praxiskonzept zu experimentieren, und die strategische Unterstützung der Organisation, die von diesem Präventionsansatz bei Gewalt in Familien überzeugt war. Im Rahmen eines Modellprojektes „Nicht wegweisen sondern einen Weg weisen“ baute die Beratungsstelle für Familien mit Gewalterfahrungen seit dem Jahr 2008 einen Kontakt zur kanadischen Programmentwicklerin Katreena Scott auf, ließ das englische Manual (Scott et al. 2006) ins Deutsche übersetzen und richtete im Jahr 2010 in Düsseldorf eine Fachtagung für die deutsche Fachpraxis und einen europäischen Austausch der Praxisanbieter*innen von Caring Dads aus. Die Durchführung einer Weiterbildung zum/zur Caring Dads Trainer:in unterstützte die Verbreitung des Ansatzes in Deutschland. Caring Dads wird seitdem zumindest an zwei weiteren Standorten in Hannover und Groß-Gerau regelmäßig durchgeführt.

Untersuchungsgegenstand

Im Rahmen der Pilotevaluation wurde das Caring Dads Programm in Düsseldorf, Hannover und Groß-Gerau untersucht. Die Standorte unterschieden sich hinsichtlich ihrer Zielgruppen und Zuweisungspraxis aufgrund unterschiedlicher institutioneller Kontexte. Das Angebot der *Fachberatungsstelle für Familien mit Gewalterfahrungen in Düsseldorf* wurde im Rahmen der Zusammenarbeit mit dem örtlichen Jugendamt schwerpunktmäßig auf diesen Überweisungskontext ausgerichtet. Die Einrichtung ist zudem in ein Interventionsprojekt gegen häusliche Gewalt eingebunden. Das *Männerbüro Hannover* bezog zwar mehrheitlich Väter durch Jugendämter, aufgrund der strukturell gewachsenen Kooperation mit der Staatsanwaltschaft im Rahmen eines Interventionsprojektes bei häuslicher Gewalt, aber auch im Rahmen der Strafverfolgung sowie durch Familiengerichte. Im *Beratungszentrum des Diakonischen Werks in Groß-Gerau* wurde Caring Dads vielmehr als selektiver Präventionsansatz der Familienbildung angeboten. Hier wurden nur Väter aufgenommen, zu denen noch keine Gewalttätigkeit in der Familie bekannt war. Die drei Standorte bildeten also das mögliche Spektrum des Caring Dads Ansatzes ab, denn alle Zielgruppen werden durch das kanadische Konzept abgedeckt (Scott et al. 2006, 2014). Die Datenerhebung erfolgte durch die Fachkräfte der Einrichtungen.

Aufbau der Untersuchung

Die Studie setzte sich aus einer Längsschnittstudie mit Befragung von teilnehmenden Vätern und den Müttern des Kindes, dem transnationalen Vergleich von teilnehmenden Vätern aus Deutschland und Kanada, einer postalischen Nachbefragung der Väter in Deutschland und einer hier nicht berichteten Stakeholder-Befragung zusammen. Methodisch erfolgte die wissenschaftliche Begleitung zur Selbstevaluation in mehreren Schritten:

1. Im Rahmen der wissenschaftlichen Begleitung des Modellprojektes bei der Diakonie Düsseldorf im Jahr 2008 wurde zunächst ein wissenschaftliches Konzept zur Selbstevaluation des Caring Dads Programms erstellt, welches eine Einordnung des Konzeptes im Rahmen der deutschen Kinder- und Jugendhilfe sowie eine Bewertung im Spiegel der internationalen Wirkungsforschung zu Elternprogrammen bei Risiken für Kindesmisshandlung und der seitens der kanadischen Programmentwicklerin *Katreena Scott* vorgeschlagenen Instrumente zur Wirkungsmessung enthielt (*Liel; Kindler 2009*).
2. Fünf Jahre später wurden aus Fallakten der Diakonie Düsseldorf die Erhebungsdaten mit diesen Instrumenten zu den ersten Programmteilnehmern ($n = 26$) in einer Machbarkeitsstudie ausgewertet (*Meindl 2014*).
3. Diese Stichprobe wurde mit weiteren Vätern aufgefüllt, die in den Jahren 2014–2016 in Düsseldorf, Hannover oder Groß-Gerau an Caring Dads teilnahmen (*Heintz-Martin; Koch 2016*). Hierfür wurden bis dato eingesetzte Instrumente zum Teil durch validierte deutsche Fassungen ersetzt, die Fachkräfte im Einsatz dieser Instrumente geschult und zur Durchführungsintegrität des Caring Dads Programms an den drei Standorten befragt.
4. Anschließend wurde diese Stichprobe mit einer Stichprobe von Caring Dads Teilnehmern aus der kanadischen Begleitforschung verglichen. Die Publikation der kanadischen Evaluationsdaten von *Scott und Lishak (2012)* bildete eine Vorlage für die deutsche Studie.
5. Im Jahr 2016 wurden alle Väter, die seit 2008 an Caring Dads teilgenommen hatten, mit einem Nachfolgezeitraum von sechs Monaten bis acht Jahren postalisch befragt (*Heintz-Martin; Koch 2016*). Die Katamnese sollte die subjektive Bewertung des Programms und der Bedeutsamkeit der Veränderungen durch die Väter erfassen. Für die Beantwortung des Fragebogens erhielten die Väter eine Incentivierung in Form eines Einkaufsgutscheins im Wert von 10 Euro.

Zur Messung von Veränderungen in den deutschen Stichproben und kulturellen Unterschieden zwischen den deutschen und kanadischen Stichproben kamen international verbreitete Verfahren zur Erfassung der selbstberichteten Änderungsmotivation, Aggressivität, dysfunktionalen Erziehungsverhaltens und Co-Parentings von Vätern sowie der väterlichen Einstellungen in der Partnerschaft und des Co-Parentings aus Müttersicht zum Einsatz (siehe Tabelle auf Seite 78). Der Katamnesebogen enthielt Fragen zur Bewertung des Programms sowie erlebter Veränderungen der Lebenssituation, Erziehungseinstellungen und familiären Beziehungen seitens der Väter.

Ergebnisse der Evaluation

In die Studie eingeschlossen wurden die Selbstberichte von 60 Vätern, die bei 45 Vätern vollständig Prä und Post vorlagen. In 34 Fällen lagen auch Berichte der Mütter vor. Die Befunde wurden an anderer Stelle bereits ausführlich berichtet (Liel et al. 2021).

Im Ergebnis zeigt sich, dass es mit dem Programm gelingt, psychosozial belastete Väter zu erreichen, denn die Väter waren häufiger arbeitslos (28 Prozent), ohne Schulabschluss (28 Prozent) oder hatten einen Migrationshintergrund (38 Prozent). Fast 40 Prozent bezogen staatliche Leistungen zur sozialen Mindestsicherung (zum Beispiel Arbeitslosengeld II). Die Teilnahme an der Maßnahme erfolgte überwiegend auf Anordnung oder Empfehlung des Jugendamtes, in über der Hälfte der Fälle wurde eine Kindeswohlgefährdung gemäß §8a SGB VIII festgestellt. Angaben von Müttern lagen bei getrennt und zusammenlebenden Elternpaaren vor, waren aber signifikant häufiger bei Kindern unter drei Jahren verglichen mit älteren Kindern.

Längsschnittliche Veränderungen sind in der Tabelle dargestellt. Mit Ausnahme bei der Änderungsmotivation wurden auf allen Skalen mittlere Verbesserungen gemessen. Die Änderungsmotivation wurde nicht als änderungssensitiv erwartet und änderte sich auch nicht. Bei signifikanten Verbesserungen wurde die Stärke des Effekts mithilfe des Vergleichsmaßes Cohens d bestimmt. Gemäß Interpretationsvorschlag können Effektstärken von $d = .20$ als klein, von $d = .50$ als moderat und von $d = .80$ als hoch bewertet werden (Cohen 1988). Bei selbstberichteter Aggressivität und dysfunktionalem Erziehungsverhalten der Väter wurden also signifikante Verbesserungen mit moderaten und hohen Effektstärken gemessen. Das elterliche Co-Parenting verbesserte sich aus Sicht der Mütter, aus Sicht der Väter jedoch nicht signifikant. Die Mütter berichteten zudem eine Abnahme feindseliger väterlicher Einstellungen in der Partnerschaft.

Bei Werten im Risikobereich der Skalen wurde darüber hinaus die Höhe der erreichten Veränderung bestimmt (klinische Signifikanz) und mittels RCI für Messfehler kontrolliert. Der RCI beruht darauf, dass die Differenz der Messwerte im Verhältnis zur Korrelation der Skalen bewertet wird. Nur bei Überschreitung eines international etablierten Grenzwertes des RCI können Messfehler sicher ausgeschlossen werden. Bei diesem Analyseschritt wurden nur Väter einbezogen, die zu T1 gemäß Grenzwert der Skalen bereits dem Risikobereich zugeordnet wurden, da nur diese Teilnehmer überhaupt die Chance für eine Verbesserung in den Normbereich hatten. Klinisch signifikante Verbesserungen vom Risiko- in den Normbereich wurden nur berücksichtigt, wenn Messfehler mit dem RCI ausgeschlossen wurden. Mithilfe dieses harten Bewertungsmaßstabs zeigten sich auf den Skalen zur Aggressivität und zum Erziehungsverhalten bei 40 Prozent und 25 Prozent der Väter, die zu Interventionsbeginn als klinisch auffällig eingestuft wurden, längsschnittliche Verbesserungen auf ein unauffälliges Niveau. Klinische Verbesserungen wurden bei Vätern kleiner Kinder (< 3 Jahre) signifikant häufiger gefunden als bei Vätern von älteren Kindern. Die Stichproben bei diesen Analysen waren allerdings klein (siehe Tabelle auf der nächsten Seite).

Konstrukt	Instrument	Quelle	Verbesserungen		Klinische Signifikanz	
			T1 vs. T2		T1 vs. T2	
			N	Cohens d ^{a,b,c}	N _{klin} ^c	Normalisiert ^d
Väter:						
Änderungsmotivation	Eigenentwicklung Caring Dads	Scott et al. (2006)	33	n.s.	-/-	-/-
Aggressivität	Aggression Questionnaire	Buss; Perry (1992)	34	.38 ^e	5	40 %
dysfunktionales Erziehungsverhalten	Parenting Scale	Arnold et al. (1993); Rhoades; O'Leary (2007)	41	.80 ^e	8	25 %
Co-Parenting	Parenting-Alliance-Measure	Abidin; Konold (1999)	37	n.s.	9	33 %
Mütter:						
väterliche Einstellungen	Quality of Life Index	Dobash et al. (2000)	18	.46 ^e	3	33 %
Co-Parenting	Parenting-Alliance-Measure	Abidin; Konold (1999)	22	.46 ^e	1	0 %

Anmerkungen:

^a Effektstärken basierend auf Mittelwertvergleichen mittels t-Test für verbundene Stichproben/Wilcoxon Signed Rank Test

^b Interpretationshilfe nach Cohen (1988): $d = .20$ schwach, $d = .50$ moderat, $d = .80$ stark

^c Teilnehmer im Risikobereich zu T1 (Basis für die RCI-Evaluation)

^d Verbesserung vom Risiko- in den Normbereich in der deutschen Stichprobe (bei signifikantem RCI)

^e Verbesserung von T1 zu T2 in der deutschen Stichprobe

Tabelle 1: In der Selbstevaluation eingesetzte Verfahren und longitudinale Ergebnisse (Liel et al. 2021)

Hinzu kamen auf allen Skalen Verbesserungen von Programmteilnehmern innerhalb eines klinisch unauffälligen Niveaus, die mittels RCI als messsicher bewertet werden konnten. Hinweise auf Wirkungen des Ansatzes wurden also sowohl zur selektiven Prävention mit belasteten, aber klinisch unauffälligen Vätern, als auch zur indizierten Prävention mit klinisch auffälligen Vätern gefunden.

In der Tabelle nicht dargestellt sind Ergebnisse des Vergleichs zwischen deutschen ($n = 60$) und kanadischen Programmteilnehmern ($n = 59$). Hier zeigten sich signifikante Mittelwertunterschiede nur auf den Skalen zur Veränderungsmotivation und auf einer Subskala des Aggression Questionnaire zur verbalen Aggression. Auf diesen Skalen berichteten deutsche Väter vor und nach dem Programm höhere Werte als kanadische Väter (Liel et al. 2021).

Im Rahmen der Katamnese mit einer Rücklaufquote von über 60 Prozent ($n = 20$) berichteten Väter in Einzelfällen von negativen Ereignissen (erneute Gewalt oder invasive Schutzmaßnahme durch das Jugendamt) und von positiven Ereignissen (erneuter Umgang oder Zusammenleben mit dem Kind) nach Abschluss des Caring Dads Programms. Die durchweg positiven Rückmeldungen zum Programm und der erlebten Veränderung zeigten die subjektive Bedeutsamkeit der durch Caring Dads vermittelten Erfahrungen und Erkenntnisse für die Väter und die hohe Relevanz des Gelernten für die Gestaltung der familiären Beziehungen.

Schlussfolgerungen

Der Mehrwert der wissenschaftlichen Begleitung zur Selbstevaluation bestand darin, dass schrittweise flexibel auf die Entwicklungen in der Praxis reagiert werden konnte. Beispielsweise war es anfänglich nicht absehbar, dass sich Caring Dads in Düsseldorf als Regelangebot festigen würde und im Rahmen einer von der Diakonie Düsseldorf angebotenen Weiterbildung weitere Standorte in Deutschland hinzukommen würden. Gleichwohl oblag die Datenerhebung den Praxiseinrichtungen, was ungleich kostengünstiger im Vergleich zu einer externen Evaluation war, aber für die Fachkräfte eine nicht unerhebliche Mehrbelastung bedeutete. Die wissenschaftlich begleitete Selbstevaluation kann somit nicht als kostengünstiger Ersatz für eine Wirksamkeitsstudie gelten, aber vorläufige Orientierung bieten.

Für die Praxis liefert die Studie die Gewissheit, konzeptionell auf dem richtigen Weg zu sein. Die Befunde deuten in Richtung des Erreichens erwünschter Wirkungen und es ergaben sich keine Hinweise auf unerwünschte Nebenwirkungen oder Problemverfestigungen. Das Caring Dads Programm scheint im selektiven und im indizierten Präventionsbereich von Kindesmisshandlung gleichermaßen geeignet zu sein, um Väter bei der Reflexion ihres Erziehungsverhaltens zu unterstützen und altersangemessenes Erziehungswissen zu vermitteln. Obwohl die Zielgruppe gewalttätiger Väter als schwer veränderbar gilt und sich in dieser Studie durchaus psychosozial belastet gezeigt hat, wurden Verbesserungen des Erziehungsverhaltens mit hohen Effektstärken gemessen. Diese können auch durch die Erhebungsmethodik „Selbstbericht“ positiv verzerrt sein. Sie halten bei einer kleinen Teilstichprobe der Kontrolle von Messerfehlern stand. Bei dieser Replikation der kanadischen Studie von Scott und Lishak (2012) fällt der in Deutschland geringere

Anteil klinisch signifikanter Verbesserungen auf. Aufgrund der geringen Stichprobengröße ist unklar, welche Verbesserungen bei dieser Zielgruppe der Sozialen Arbeit überhaupt möglich sind.

Bei den gemessenen Unterschieden zwischen Vätern in Deutschland und Kanada handelt es sich um reale Unterschiede. In Deutschland wird Caring Dads häufiger in der Kinder- und Jugendhilfe genutzt, während es in Kanada überwiegend unter Auflagen im Rahmen der Strafverfolgung eingesetzt wird. Eine höhere Eigenmotivation der Programmteilnehmer in Deutschland erklärt sich mit dieser Zuweisungspraxis. Unterschiede bei verbaler Aggressivität deuten eher auf die Aufdeckung eines realen kulturellen Unterschiedes im Sprachgebrauch hin, als dass sie die Übertragbarkeit des Ansatzes nach Deutschland infrage stellen würden.

Die Erkenntnisse sind methodisch das Verdienst der wissenschaftlichen Begleitung, da die Datenauswertung die Möglichkeiten der Selbstevaluation der Praxis überstieg. Das Vorgehen kann dabei helfen, zum Zeitpunkt der praktischen Erprobung eines Ansatzes Wirkungswissen zu generieren. Es ist damit ein Ausweg, um den gängigen Schwierigkeiten der Wirkungsforschung in der Sozialen Arbeit vorläufig zu begegnen (Liel 2018).

Bezogen auf das Caring Dads Programm belegt eine aktuelle quasi-experimentelle Interventionsstudie aus Kanada auch Verringerungen von Kindesmisshandlung (Scott et al. 2021). Solche methodisch ungleich anspruchsvolleren Befunde können auf der Grundlage eines international vergleichenden Evaluationsansatzes – wie hier praktiziert – besser genutzt werden, da die Wirkweise zwischen beiden Ländern bereits verglichen wurde.

Die Studie ist auch ein Beispiel der Nutzung von diagnostischen Erhebungsdaten zum Zwecke der Wirkungsforschung. Solche Nutzungskonzepte dürften mit der steigenden Digitalisierung der Dokumentationspraxis in Einrichtungen der Sozialen Arbeit zukünftig wahrscheinlich leichter fallen.

Aus wissenschaftlicher Sicht beinhaltet dieses Vorgehen methodische Kompromisse, zum Beispiel hinsichtlich der Datenqualität und Durchführungsintegrität, um den praxisbezogenen Evaluationsansatz in dieser Form umsetzen zu können. Mithilfe der Konzepte der klinischen Signifikanz und des RCI konnte die interne Validität der Studie zumindest verbessert werden. Bei eingeschränkter externer Validität konnten zu einem frühen Zeitpunkt der Implementierung für die Praxis der Sozialen Arbeit bedeutsame Befunde generiert werden. Diese Erkenntnisse können als Grundlage für die weitere Forschung zur Arbeit mit Vätern zur Gewaltprävention und im Kinderschutz dienen und möglicherweise ein Schritt zu mehr Wirkungsorientierung in der Sozialen Arbeit in Deutschland sein.

Dr. phil. Christoph Liel, M.A., Dipl.-Sozialarbeiter, ist wissenschaftlicher Referent und Projektleiter im Nationalen Zentrum Frühe Hilfen am Deutschen Jugendinstitut e.V. in München.

Doreen Herler, Dipl.-Pädagogin, ist stellvertretende Geschäftsführerin im Männerbüro Hannover und Fachkraft für Caring Dads.

Edgar Schulz, Lehrer i. R., war Leiter des Arbeitsgebiets Erwachsenengruppen und Caring Dads Pionier in der Fachberatungsstelle für Familien mit Gewalterfahrungen der Diakonie Düsseldorf.

Literatur

- Abidin, Richard R.; Konold, Timothy R. (1999): PAM Parenting Alliance Measure. Professional manual. Charlottesville, Virginia
- Arnold, David S.; O'Leary, Susan G.; Wolff, Lisa S.; Acker, Maureen M. (1993): The Parenting Scale. A measure of dysfunctional parenting in discipline situations. In: *Psychological Assessment* 2/1993, S. 137-144
- Borrmann, Stefan; Thiessen, Barbara (Hrsg.) (2016): *Wirkungen Sozialer Arbeit. Potentiale und Grenzen der Evidenzbasierung für Profession und Disziplin*. Opladen
- Buss, Arnold H.; Perry, Mark (1992): The Aggression Questionnaire. In: *Journal of Personality and Social Psychology* 3/1992, S. 452-459
- Cohen, Jacob (1988): *Statistical power analysis for the behavioral sciences*. Hoboken
- Dobash, R. Emerson; Dobash, Russel P.; Cavanagh, Kate; Lewis, Ruth (2000): Changes in quality of life. In: Dobash, R. Emerson; Dobash, Russel P.; Cavanagh, Kate; Lewis, Ruth (Hrsg.): *Changing violent men*. Thousand Oaks, S. 127-146
- Dudenredaktion (2021): Wirkung. In: <https://www.duden.de/node/206276/revision/206312> (veröffentlicht o.J., abgerufen am 03.11.2021)
- Engelke, Ernst; Spatscheck, Christian; Borrmann, Stefan (2016): *Die Wissenschaft Soziale Arbeit. Werdegang und Grundlagen*. Freiburg im Breisgau
- Heintz-Martin, Valerie; Koch, Marlene (2016): *Wirkung des Caring Dads Programms zur Prävention von Kindesmisshandlung und Vernachlässigung mit Vätern in Deutschland. Abschlussbericht für das Deutsche Jugendinstitut*. München
- Jacobson, Neil S.; Truax, Paula (1991): Clinical significance. A statistical approach to defining meaningful change in psychotherapy research. In: *Journal of Consulting and Clinical Psychology* 1/1991, S. 12-19
- König, Joachim (2007): *Einführung in die Selbstevaluation. Ein Leitfaden zur Bewertung der Praxis Sozialer Arbeit*. Freiburg im Breisgau
- Liel, Christoph (2017): Täterarbeit bei Partnergewalt. Auswirkungen auf das Rückfallrisiko. In: *Forensische Psychiatrie, Psychologie und Kriminologie* 1/2017, S. 59-68
- Liel, Christoph (2018): *Väter und familiäre Gewalt*. München
- Liel, Christoph; Kindler, Heinz (2009): *Selbstevaluation des Caring Dads Programms. Wissenschaftliches Konzept*. München
- Liel, Christoph; Koch, Marlene; Eickhorst, Andreas (2021): Arbeit mit Vätern zur Prävention von Kindesmisshandlung. Eine Pilotevaluation des Caring Dads Programms in Deutschland. In: *Praxis der Kinderpsychologie und Kinderpsychiatrie* 2/2021, S. 115-133
- McConnell, Nicola; Barnard, Matt; Taylor, Julie (2017): Caring Dads Safer Children. Families' perspectives on an intervention for maltreating fathers. In: *Psychology of Violence* 3/2017, S. 406-416
- Meindl, Lisa (2014): *Evaluation eines Präventionsprogramms bei Risiken von Kindeswohlgefährdung und Kindesmisshandlung mit Vätern*. Masterthesis. Ludwig-Maximilians-Universität. München

- Rhoades, Kimberly A.; O'Leary, Susan G. (2007): Factor structure and validity of the parenting scale. In: *Journal of Clinical Child & Adolescent Psychology* 2/2007, S. 137-146
- Scott, Katreena; Dubov, Violeta; Devine, Christine; Colquhoun, Chrystal; Hoffelner, Carrie; Niki, Izumi; Webb, Sarah; Goodman, Deborah (2021): Caring Dads intervention for fathers who have perpetrated abuse within their families. Quasi-experimental evaluation of child protection outcomes over two years. In: *Child Abuse & Neglect*, S. 105-204
- Scott, Katreena; Francis, Karen; Crooks, Claire; Kelly, Tim (2006): *Caring dads. Helping fathers value their children*. Victoria, B.C
- Scott, Katreena L.; Kelly, Tim; Crooks, Claire; Francis, Karen (2014): *Caring dads. Helping fathers value their children*. North Charleston
- Scott, Katreena L.; Lishak, Vicky (2012): Intervention for maltreating fathers. Statistically and clinically significant change. In: *Child Abuse & Neglect* 9/2012, S. 680-684
- Sommerfeld, Peter; Hüttemann, Matthias (Hrsg.) (2007): *Evidenzbasierte soziale Arbeit. Nutzung von Forschung in der Praxis*. Baltmannsweiler
- Suess, Gerhard J.; Bohlen, Uta; Mali, Agnes; Maier, M. Frumentia (2010): Erste Ergebnisse zur Wirksamkeit Früher Hilfen aus dem STEEP-Praxisforschungsprojekt „WiEge“. In: *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz* 11/2010, S. 1143-1149
- Zimmermann, Peter: Vorhandene Maßnahmen und Programme mit Vätern und deren Wirkung. In: Eickhorst, Andreas; Liel, Christoph (Hrsg.): *Väter in den Frühen Hilfen. Ein Sammelband für Wissenschaft und Praxis*. Köln im Druck

Sigrid James, Franziska Seidel, Julian Trostmann & Terhas Andezion

Was ist Gelingen? Erkenntnisse eines Modellprojekts zur Arbeitsmarktintegration von Frauen mit Migrationshintergrund

Einleitung

Menschen mit Migrationshintergrund, insbesondere Frauen, sind bei der Arbeitsmarktintegration mit unterschiedlichen Hürden konfrontiert und erfahren deutliche Benachteiligungen.

¹ Als Teil spezifischer Integrationsförderangebote auf Bundes- sowie Landesebene sind daher in den letzten Jahren eine Vielzahl an Projekten generell und für Frauen im Besonderen initiiert worden (für eine erste Übersicht siehe Die Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration, 2021). Wirkungsorientierte Evaluationen dieser Initiativen bleiben bisher eine Ausnahme, aber im Gesamtbild der verfügbaren evaluativen Studien werden positive Ergebnisse in Bezug auf diverse Aspekte der Arbeitsmarktintegration verzeichnet (für einen Überblick sowie einzelne Studien siehe u. a. Aumüller, 2016, S. 37; Dobischat & Schäfer, 2020; James et al., 2020; Schäfer & Dobischat, 2021; Walter et al. 2014). Zudem sind Erkenntnisse über mögliche Wirkfaktoren generiert worden, die die konzeptionelle Grundlage für neuere Projekte in diesem Bereich bilden.

Arbeitsmarktintegrationsprojekte stellen die Evaluationsforschung vor eine große Herausforderung. Sie untersuchen Projekte, die an der Schnittstelle unterschiedlicher politischer, projektbezogener sowie individueller Interessen und Zielvorstellungen stehen, in komplexe Strukturen eingebunden sind und von vielfachen organisationsinternen sowie externen Faktoren beeinflusst werden. Dies bedeutet eine ungünstige Ausgangslage, um aus wissenschaftlicher Sicht wirkmächtige Beziehungen zu untersuchen. Fehlende Vergleichsgruppen, diffuse Erfolgsindikatoren, heterogene Zielgruppen sowie dynamische Projektkonzepte bilden zudem Barrieren für auf Wirkung orientierte Evaluationsstudien (Bleck, 2016, S. 111; Merchel, 2019, S. 138 f.). Wie bei der Teilnahme an einem Arbeitsmarktintegrationsprojekt Gelingen vonseiten unterschiedlicher Beteiligter definiert wird, ist für ein Verständnis und die Darstellung des Erfolgs solcher Projekte jedoch ebenso von Relevanz, wie auch die Frage nach der Wirksamkeit der angebotenen Unterstützungsleistungen.

Nicht zuletzt vor diesem Hintergrund sind Probleme dabei zu benennen, empirisch gewonnenes Wissen mit Persistenz und Nachhaltigkeit in die Handlungspraxis zu transferieren (Gollwitzer & Jäger, 2014, S. 113). Dieser Wissenstransfer stößt immer wieder an Grenzen, wenn theoretische 'good' oder 'best practice' auf Dynamiken und Heterogenitäten in der Handlungspraxis stößt (zu den allgemeinen Problemen des Wissenstransfers vgl. u. a. Oestreicher, 2014). Die beschriebenen

¹ In diesem Beitrag wird der Begriff „Frauen mit Migrationshintergrund“ genutzt. Der Begriff an sich wird kontrovers diskutiert, und es gibt Vorschläge von der Fachkommission Integrationsfähigkeit (2021, S. 10) für treffendere Beschreibungen. Da bisher kein Begriff in diesem Bereich unproblematisch ist, wird in diesem Beitrag weiterhin „Frauen mit Migrationshintergrund“ verwendet. Der Begriff reflektiert zudem die Sprachlichkeit des Projekts und erfasst die Heterogenität dieser Gruppe, die Frauen mit Fluchterfahrungen sowie Frauen, die selbst oder deren Familien nach Deutschland zugewandert sind, umschließt. Rund die Hälfte (52%) der in Deutschland lebenden Personen mit Migrationshintergrund besitzt die deutsche Staatsangehörigkeit (Statistisches Bundesamt, 2019, n.d).

Komplexitäten attestieren die Notwendigkeit von Forschungsmethoden, die umfassend, iterativ, flexibel und kreativ sind (James, 2019, S. 47). Von Interesse ist weniger oder nicht nur die Frage, was wirkt, sondern auch, was wie für wen unter welchen Bedingungen wirkt (u. a. Nüsken, 2015, S. 20; Ottmann & König, 2018, S. 5 f.). Es geht also in erster Linie nicht darum, Störfaktoren, die die interne Validität einer Studie gefährden, zu eliminieren, sondern sie als kontextuelle Faktoren zu verstehen und die Prozesse und Dynamiken zu erfassen, die einem Projekt unterliegen. In diesem Zuge haben Mixed Method Evaluationen sowie Realistische Evaluationen an Bedeutung in der Sozialen Arbeit gewonnen (Bleck, 2016, S. 111). Aufgrund methodischer Herausforderungen, ethischer Bedenken sowie der nachweislichen Schwierigkeiten experimentell begründete Programme und Handlungsmethoden in die aktuelle Praxiswelt zu transferieren, bleibt daher die randomisierte Kontrollgruppenstudie, die immer noch als Goldstandard für die Bestimmung von Wirkung gilt, das am wenigsten genutzte Forschungsdesign in der sozialarbeitswissenschaftlichen Evaluations- und Wirkungsforschung (Holosko, 2010, S. 665).²

Anhand des Beispiels eines Kassler Modellprojekts zur Arbeitsmarktintegration von Frauen mit Migrationshintergrund und seiner wissenschaftlichen (primär formativen) Begleitstudie beschreibt dieser Beitrag, wie sich trotz der genannten Barrieren mit Fragen des Gelingens auseinandergesetzt werden kann. Zu diesem Zweck wird zuerst ein kurzer Überblick über Inhalte der Arbeitsmarktintegration von Frauen mit Migrationshintergrund gegeben, das Projektkonzept vorgestellt und die Begleitstudie in ihrem methodischen Ansatz skizziert. Darauf folgt eine Betrachtung unterschiedlicher Perspektiven des Gelingens sowie eine durch konkrete Beispiele untermauerte Darstellung methodischer Überlegungen und Lösungsversuche, die darauf abzielen, eine Balance zwischen forschungsmethodischer Rigorosität und praxisnaher Relevanz zu erreichen.

Arbeitsmarktintegration von Frauen mit Migrationshintergrund in Deutschland

Frauen mit Migrationshintergrund sind auf dem deutschen Arbeitsmarkt nachgewiesen benachteiligt (Sachverständigenrat Deutscher Stiftungen für Integration und Migration (SVR), 2019, S. 113; Salikutluk et al., 2020, S. 19). Sie verdienen im Durchschnitt weniger (Lukas, 2011, S. 18; SVR, 2019, S. 116), haben einen beruflichen Status, der häufig nicht ihrer eigentlichen Qualifikation entspricht (Ferenschild, 2019, S. 8) und sind mit Vorurteilen und Diskriminierung konfrontiert (Antidiskriminierungsstelle des Bundes, 2020, S. 15; Farrokhzad, 2006, S. 64). Mangelnde Sprachkenntnisse werden oft mit fehlender Fachkompetenz gleichgesetzt (Esser, 2006, S. 12, 89) und bürokratische Hindernisse sowie Probleme bei Anerkennungsverfahren der im Heimatland erworbenen Berufs- oder Bildungsabschlüsse erschweren den Zugang zum Ausbildungs- und Arbeitsmarkt in Deutschland (Ferenschild, 2019, S. 9; Salikutluk et al., 2020, S. 27). Unzureichende außerfamiliäre Kinderbetreuungsmöglichkeiten sind ein Faktor, der die Teilnahme an Arbeitsmarktintegrationsprojekten für Frauen mit Migrationshintergrund besonders erschwert (Farrokhzad, 2018, S. 170). Die vielfachen Herausforderungen tragen zu psychosozialen Belastungen, wie sozialer Isolation und erhöhtem Stresserleben, bei – Faktoren, die den Zugang zum Arbeitsmarkt zusätzlich erschweren und auch ein Hindernis für die erfolgreiche Teilnahme

² Für ein differenziertes Verständnis von 'Wirkung' und ihrer Messung und Analyse siehe Ottmann und König (2018).

an Arbeitsmarktintegrationsprojekten sein können (Aichberger et al., 2015, S. 1697; Irfaeya, et al., 2008, S. 342 f; Kluge et al., 2019, S. 1103).

Vor dem Hintergrund der genannten Problemlagen und dem generellen Bemühen, Ressourcen von Frauen mit Migrationshintergrund zu stärken, rückt diese Zielgruppe seit einigen Jahren vermehrt in den Fokus von Angeboten zur Arbeitsmarktintegration (Die Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration, 2019, S. 175 f.; Die Bundesregierung, 2007, S. 61, 95). Eine Zunahme der frauenspezifischen Förder- und Integrationsangebote in den letzten Jahren muss nicht zuletzt auch im Kontext der Fachkräftestrategie der Bundesregierung und dem vorhandenen Fachkräftemangel, insbesondere in sozialwirtschaftlichen und pflegerischen Berufen, gesehen werden (Bundesgesundheitsministerium, 2021, n.d.). Die diversen Arbeitsmarktintegrationsprojekte, auch für diese Zielgruppe, beruhen häufig auf ähnlichen konzeptuellen Ansätzen und umfassen zumeist Projektbausteine wie beispielsweise Coaching, Unterstützung bei der Kinderbetreuung sowie Sprachförderung (u. a. Aumüller, 2016, S. 37; BMFSFJ, 2018, S. 12; 2019, S. 5). Wie schon angedeutet, bleibt die empirische Datenlage in Bezug auf Frauen mit Migrationshintergrund trotz gestiegenem Interesse weiterhin lückenhaft. Ebenso gehören Untersuchungen zur Wirksamkeit der Angebote sozialer Dienste und insbesondere auch der Arbeitsmarktintegrationsprojekte in Deutschland noch nicht zum Standard (Die Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration, 2019, S.143, 176; Habermann & Stagge, 2015, S. 173). Vor dem Hintergrund, dass die Soziale Arbeit am Theorie-Praxis-Wissenstransfer und nachhaltigen Lösungen für Adressat*innen interessiert ist und zudem zunehmend aufgefordert wird, ihre Maßnahmen und Hilfeleistungen auch empirisch zu untermauern, erscheint es notwendig, verstärkt die Qualität und Wirksamkeit bestehender Angebote zu bedenken und zu untersuchen.

Das Modellprojekt „Neue Chancen in der Sozialwirtschaft – Qualifizierungsperspektive für Migrantinnen“

Das 2018 initiierte Modellprojekt „Neue Chancen in der Sozialwirtschaft – Qualifizierungsperspektive für Migrantinnen“ der Stadt Kassel (Abk. SoWi) ist ein bis 2022 laufendes Arbeitsmarktintegrationsprojekt, das durch das Förderprogramm des Hessischen Ministeriums für Soziales und Integration „Sozialwirtschaft integriert“ unterstützt und von der Kommunalen Arbeitsförderung im Sozialamt der Stadt Kassel umgesetzt wird. Zielgruppe für das Projekt sind gewöhnlich Frauen mit Migrationshintergrund zwischen 18 und 45 Jahren, die einen (Wieder-)Einstieg in das Berufsleben anstreben. Neben der Eröffnung von Arbeitsmarktperspektiven sind die formale Weiterqualifizierung (d. h., Teilnahme an einem Hauptschulkurs, Sprachkursen oder Praktika) und die Qualifizierung in sozialwirtschaftlichen Berufen (insb. Erziehung, Pflege, Hauswirtschaft) Kernelemente des Projekts. Explizites von der Politik gestecktes Ziel ist ein qualifizierter Berufsabschluss, der von (mindestens) 50 Frauen in der Laufzeit des Projekts erreicht werden soll.

Als Eingangsvoraussetzung für die Projektteilnahme gilt i. d. R. ein Nachweis von B1 Deutschkenntnissen (vgl. GER) sowie ein ausreichendes Maß an Lern- und Schulerfahrung, das in einem sozialanamnetischen Gespräch ermittelt wird. Der Einstieg beginnt für die meisten Frauen mit

einer dreimonatigen Orientierungsphase, in der berufliche Ziele mit der Eignung der Frauen abgeglichen, unterschiedliche Inhalte zu relevanten Berufsfeldern und Themen vermittelt sowie ein Praktikum absolviert werden. Nach dieser Orientierungsphase, die in Kohorten von ca. 20 Frauen stattfindet, werden die Teilnehmerinnen individuell, primär durch Fachkräfte der Sozialen Arbeit unterschiedlicher kooperierender Träger, über mehrere Monate gecoacht. Dieses Coaching ist ein Kernelement des Projekts, das einen gelingenden Qualifikationsprozess der Teilnehmerinnen gewährleisten soll (Hessisches Ministerium für Soziales und Integration, n.d., S. 8 f.). Die individuelle Betreuung ist genderspezifisch und kann von den Coachinnen je nach den Bedürfnissen der Teilnehmerinnen ausgestaltet und angepasst werden, d. h. das Projektkonzept wird als dynamisch und offen verstanden sowie stetig in Zusammenarbeit mit den Kooperationspartnerschaften weiterentwickelt.

Bis Februar 2021 wurden 192 Projektaufnahmen verzeichnet. Daten, die in diesem Beitrag illustrierend hinzugezogen werden, beziehen sich auf Zwischenberichtsauswertungen vom Sommer und Herbst 2020 – ca. zwei Jahre nach Projektbeginn.

Die begleitende Evaluationsstudie

Das vierjährige Modellprojekt wird seit Beginn des Projekts wissenschaftlich begleitet. Die Herausforderungen für die Begleitstudie, in der eine empirisch fundierte 'Außensicht' erfasst werden soll, sollten aufgrund der Projektbeschreibung offensichtlich sein: 1) SoWi ist ein neues Projekt, das insbesondere in der Anfangsphase viele Veränderungen durchlaufen hat; 2) die Merkmale des Projekts (dynamisch, individualisiert, offen, etc.) erschweren eine auf Wirkung orientierte Evaluation; 3) Kernelemente des Projekts werden von Kooperationspartnerschaften umgesetzt; 4) es gibt bisher keine Leitfäden; 5) Monitoring Daten, die essentiell für die Beschreibung von Gelingen sind, werden nicht vom Evaluationsteam erfasst, sondern beruhen insbesondere auf Berichten der Coachinnen. Aus diesen Gründen war eine primär wirkungsorientierte Evaluation von vorneherein ausgeschlossen worden.

Zielsetzungen der Evaluation wurden in Absprache mit der Leitung der Kommunalen Arbeitsförderung der Stadt Kassel in der Planungsphase des Projekts spezifiziert und umfassen zum einen die Untersuchung von Prozessen, denen das Projekt unterliegt. Zum anderen geht die Evaluation konkreten Fragestellungen nach, die zuvörderst darauf abzielen, die Lebenslagen der Frauen zu untersuchen, psychosoziale Ressourcen und Barrieren zu identifizieren, ihre Wahrnehmung in Bezug auf akkultorative Prozesse zu erfassen und somit Profile der Frauen und ihren Erwartungen für das Projekt zu erstellen. Dies beinhaltet auch eine Untersuchung des prädiktiven Einflusses der personen- sowie umfeldbezogenen Ressourcen der Teilnehmerinnen auf die Zielgrößen des Projekts. Ferner soll die Rolle der Coachinnen und ihre Bedeutung in Bezug auf die Ziele des Projekts im Kontext möglicher Faktoren exploriert werden.

Die Evaluation erhebt über den Zeitraum des Projekts hinweg quantitative sowie qualitative Daten aus unterschiedlichen Quellen. Tabelle 1 verschafft einen Überblick über verfügbare Datenquellen zum Zeitpunkt des Zwischenberichts. Weitere Datenerhebungen haben seitdem stattgefunden und werden bis Ende 2021 durchgeführt. Forschungsmethodische Einzelheiten (u. a. Rekrutierung der

Teilnehmerinnen, Datenerhebung, Instrumentarien) werden an dieser Stelle nicht vertieft. Der Tabelle 1 können die thematischen Bereiche und einbezogenen Konstrukte entnommen werden, die, wenn vorhanden, mit elaborierten Skalen/Inventaren erhoben wurden. Grundlage der illustrierenden Ausführungen in diesem Beitrag sind die Erstbefragungen der Teilnehmerinnen, qualitative Interviews mit Expertinnen sowie Monitoring Daten.

Datenquelle	N	Erhoben von	Erfasste Bereiche
Teilnehmerinnen Erstbefragung	82	Evaluationsteam	Stressbelastung, Resilienz, soziale Unterstützung, demographische Daten, Akkulturation, etc.
Berichte der Berufsorientierung	75	Kooperationspartnerschaften	Berufswünsche, Stärken, Fortschritt im Projekt, etc.
Qualitative Expertinnen Interviews	3	Evaluationsteam	Projektkonzept, Projektumsetzung, Barrieren, Zusammenarbeit mit Akteur*innen
Qualitative Interviews mit Teilnehmerinnen	1	Evaluationsteam	Erleben im Projekt, Herausforderungen, Bewältigungsstrategien, etc.
Online Erstbefragung der Coachinnen	27	Evaluationsteam	Rollenverständnis, Barrieren, Hilfebeziehung, etc.
Monitoring Daten	152	Coachinnen / SoWi-Team	Soziodemographische Daten, Status im Projekt

Tabelle 1: Datenquellen – Stand Sommer/Herbst 2020

Was ist Gelingen? Perspektiven unterschiedlicher Akteur*innen

Erfolg kann, wird Kränzl-Nagl, Lehner und Prinz (2019) gefolgt, auch als „anerkannte Wirkung der Maßnahme, als Nutzen für Einzelne, Zielgruppen und die Gesellschaft“ (S. 23) verstanden werden und ist immer von definierten Indikatoren, bzw. Zielen, abhängig. Auch wenn die Begleitstudie nicht als Wirkungsstudie ausgelegt ist, sind Fragen des Gelingens für dieses Projekt zentral. In der wissenschaftlichen Begleitung und engen Zusammenarbeit zwischen Projekt- und Evaluationsteam werden fortwährend Erkenntnisse ausgetauscht und Status-Updates kommuniziert, auch um die Chancen einer Verstetigung zu erhöhen. Reliable Daten bilden die Grundlage, um den Verlauf und Status der Teilnehmerinnen zu unterschiedlichen Zeitpunkten im Projekt beschreiben zu können. Wie die Perspektiven, Zielsetzungen und Erfolgsstandards der diversen Akteur*innengruppen zu unterschiedlichen Definitionen des Gelingens beitragen, wird im Folgenden beschrieben.

Gelingen aus Sicht der Politik

Das Projekt SoWi wurde 2018 als Modellprojekt von kommunalpolitischen Vertreter*innen initiiert und durch die Förderung des Hessischen Ministeriums für Soziales und Integration und der Stadt Kassel umgesetzt. Die politische Initiative ist im Kontext demographischer Entwicklungen sowie der sogenannten Fachkräftestrategie zu verstehen, durch die inländische, europäische und außereuropäische Fachkräftepotenziale insbesondere in sozialwirtschaftlichen und pflegerischen Berufsfeldern gefördert werden sollen (Die Bundesregierung, 2020, n.d.; Grünheid & Sulak, 2016, S. 6 f.). Menschen mit Migrationshintergrund sind eine explizite Zielgruppe dieser

Fachkräftestrategie. Gleichzeitig verfolgt die deutsche Integrationspolitik auf Bundes- sowie Länderebene Ziele der beruflichen und gesellschaftlichen Integration von Menschen mit Migrationshintergrund (Die Bundesregierung, 2007, S. 13, 173), was insbesondere die Förderung von Frauen mit Migrationshintergrund umfasst (Hessisches Ministerium für Soziales und Integration, 2018, S. 56, 71.). Hierfür gibt es diverse Förderlinien, unter ihnen auch „Sozialwirtschaft integriert“, über welches das hier beschriebene Modellprojekt der Stadt Kassel ko-finanziert wird. Politisches Ziel für SoWi ist die Qualifizierung von mindestens 50 Frauen mit Migrationshintergrund für den deutschen Ausbildungs- und Arbeitsmarkt und die zählbare Vermittlung in eben jenen, sowie die Senkung sozialstaatlicher Hilfeleistungen durch finanzielle Unterstützung von Arbeitsmarktprojekten. Zugleich soll das Projekt einen Beitrag zur Fachkräftegenerierung für den gesundheits- und sozialpflegerischen Bereich leisten.

Gelingen aus Sicht der Projektpraxis

In erster Linie orientiert sich das Projekt an den Vorgaben, die vonseiten der Politik gestellt werden. Das heißt, Kernziel für das SoWi-Projektteam ist es, 50 Frauen im Bereich der Sozialwirtschaft in die Berufsausbildung und anschließend in die Erwerbsarbeit zu begleiten. Dies wurde als Maßstab für den Erfolg des Projekts auch in Expertinneninterviews mit Mitgliedern des Projektteams bestätigt. Die Übergangsbegleitung wird dabei verstanden als Weg in größere Unabhängigkeit und als Mittel sowie Indikator für Integration und gesellschaftliche Teilhabe. Die Förderung von Unabhängigkeit bezieht sich dabei nicht nur auf wirtschaftliche Überlegungen, sondern beinhaltet explizit die Stärkung von Frauen in ihren individuellen Lebenslagen. Aus Sicht des Projektteams ist daher eine individuelle, bedarfs- und ressourcenorientierte Förderung der Teilnehmerinnen ein zentrales Element in der Umsetzung des Projektkonzeptes. Die Stärkung von Resilienz, eine erhöhte Selbstwirksamkeit sowie größeres Selbstvertrauen werden wiederholt vom Projektteam als sekundäre Indikatoren des Gelingens erwähnt. Demnach kann Erfolg auch bedeuten, dass die teilnehmenden Frauen sich ihren eigenen Wünschen und Fähigkeiten bewusst(er) werden und diese weiterentwickeln können, selbst wenn sie das Ziel der arbeitsmarktlichen Eingliederung vorerst nicht erreichen sollten. Aus Sicht des Projektteams wird SoWi jetzt schon als Erfolg wahrgenommen, nicht zuletzt aufgrund der offenkundigen Erfolge einzelner Frauen und ihren z. T. beeindruckenden sowie bewegenden Biographien.

Gelingen aus Sicht der Teilnehmerinnen

Zentral bei der Frage des Gelingens des Projekts ist die Auseinandersetzung mit den persönlichen und berufsbezogenen Erwartungen der Teilnehmerinnen. Daten aus der Erstbefragung geben Hinweise darauf, welche Zielsetzungen Frauen mit ihrer Teilnahme am Projekt verfolgen. Eine Ausbildung zu beenden und Erwerbsarbeit zu finden waren die am meisten benannten Ziele, gefolgt von Erwartungen für die persönliche Entwicklung. Eine Reihe weiterer Zielsetzungen wurden zwar benannt, aber die primären Zielsetzungen konvergieren mit den explizierten Erwartungen und Zielen für das Projekt. Konvergenz in den Zielsetzungen wird als bedeutsamer Gelingensfaktor in Arbeitsmarktintegrationsprojekten benannt (Aumüller, 2016; S. 49 f.; James et al., 2020, S. 558). Wie Frauen ihre Teilnahme an SoWi erleben und wie sie für sich Erfolg im Projekt definieren, sind Fragen, die in qualitativen Interviews weiter vertieft werden sollen.

Gelingen aus Sicht der Wissenschaft

Was eine wissenschaftliche Begleitung wie diese grundlegend von der experimentellen Wirkungsforschung unterscheidet, ist, dass sie sich nicht auf die 'Außensicht' beschränkt, sondern durch den engen Austausch mit dem Projektteam sowie beobachtender Teilnahme an Projektphasen Einfluss auf die Projektentwicklung nimmt; d. h. Prozesse werden aufgrund empirischer Daten mitgestaltet. Dieser formative Prozess ist intendiert und unterstützt eine empirisch fundierte Weiterentwicklung des Projekts, die emergente Erkenntnisse über zentrale Gelingensfaktoren und Wirkmechanismen sowie Implementierungsbarrieren, wo immer möglich, noch in der aktiven Projektphase integriert. Er bedeutet jedoch auch, dass Fragen der Wirkung nur bedingt beantwortet werden können.³ Trotz allem werden Möglichkeiten gesehen, sich auch innerhalb einer primär formativ konzipierten Evaluation mit der Gelingensfrage methodisch rigoros auseinanderzusetzen und eine Grundlage dafür zu schaffen, sich Fragen des Gelingens durch methodische Verbesserungen anzunähern.

Methodische Überlegungen und Lösungsansätze

Im abschließenden Teil dieses Beitrags werden methodische Lösungsvorschläge im Hinblick auf selekte wissenschaftliche Aspekte vorgestellt, die für Fragen des Gelingens Relevanz haben und von Bedeutung in dynamischen Projekten wie SoWi sind.

Gesamtgröße erfassen

Trotz der differenzierten Zielsetzungen, die unterschiedliche Akteur*innengruppen verfolgen, ist die Zahl 50 in Bezug auf Teilnehmerinnen, die einen qualifizierten Berufsabschluss in der Projektlaufzeit erwerben, ein erster 'harter' Maßstab für SoWi, der durch Monitoring-Daten zu Ende des Projekts berechnet werden kann. Unklarer ist, wie viele Frauen in das Projekt eintreten können, d. h. eine Erfassung der Gesamtgröße ist entscheidend, um der Zahl 50 Bedeutung zu geben. Zu Beginn war das Projekt vonseiten der Projektplaner*innen für 120 Frauen vorgesehen, was impliziert, dass ein Erreichen der Zielsetzung von ca. 40% der Teilnehmerinnen als Erfolg beschrieben werden kann. Inzwischen ist das Projekt aufgrund diverser Faktoren (u. a. starke Nachfrage, personelle Erweiterungen) nach nur zweieinhalb Jahren von nahezu 200 Frauen begonnen worden. Die aktuelle Anzahl von Frauen im Projekt ist eine äußerst dynamische Zielgröße, die sich mit dem Ein- und Ausstieg von Frauen stetig verändert: Es gibt neue Teilnehmerinnen; Frauen, die in einer Kohorte nachrücken; Teilnehmerinnen, die aufgrund von veränderten Lebensumständen im Projekt pausieren; Frauen, die frühzeitig aus dem Projekt ausscheiden sowie eine zunehmende Anzahl an Frauen, die das Projekt erfolgreich beendet haben, also einen qualifizierenden Berufsabschluss erworben haben.

³ Aussagen über Wirkung beruhen in erster Linie auf der internen Validität des Forschungsdesigns, womit die randomisierte Kontrollgruppenstudie als der Goldstandard gilt, um die Nettowirkung einer Intervention unter Ausschluss konfundierender Einflüsse zu determinieren (Schneider, 2016, S. 204 f.). Weitere bedeutsame Kriterien, die in der Bewertung von Wirkungsstudien eine Rolle spielen, beinhalten messbare Erfolgsindikatoren, die Nutzung valider und reliabler Instrumentarien, homogene Untersuchungsgruppen, kontrollierte Datenerhebungsabläufe sowie manualisierte Projektkonzepte (Gollwitzer & Jäger, 2014, S. 66 f., 112 f.).

Gelingen relativieren

Eine Zwischenanalyse zur Projekthalbzeit erfasste den Status der an der Evaluation teilnehmenden Frauen und bezeugte die Notwendigkeit einer nuancierten Konzipierung von Erfolgsindikatoren. Laut dieser Daten erhielten 61,0% der an der Evaluation teilnehmenden Frauen⁴ ein Coaching, was einen aktiven Status im Projekt impliziert; 39,0% der Teilnehmerinnen hatten das Projekt abgebrochen. Eine binäre Konzipierung – 'aktive Teilnehmerinnen/frühzeitige Abbrüche' – erscheint durchaus plausibel um Gelingen zu erfassen und bietet Möglichkeiten für statistische Auswertungen, übersieht jedoch wichtige Nuancen. So können Frauen, die im Coaching sind, einen unterschiedlichen Status im Projekt haben. Einige erhielten ausbildungsvorbereitende Unterstützung durch Hauptschul- oder Sprachkurse, andere standen schon in der berufsqualifizierenden Ausbildung, hatten sie bereits abgeschlossen oder sogar ein Arbeitsverhältnis begonnen. Daten wiesen ferner darauf hin, dass 16% der Frauen in einem Übergangsstadium standen und ausschließlich ein Coaching erhielten. Diese Option wurde durch die Daten offensichtlich und veranlasste Überlegungen vonseiten des SoWi-Teams, inwieweit und für wie lange dieser 'Coaching-only'-Status im Projekt möglich sein sollte.

Ebenso wie aktive Teilnahme am Projekt differenziert verstanden werden muss, ist ein frühzeitiges Ausscheiden auch nicht automatisch als Misserfolg zu bewerten. Einige Frauen wechselten den Wohnort, gingen in den Mutterschaftsurlaub oder fanden eine sozialversicherungspflichtige Beschäftigung außerhalb des SoWi-Projekts. Als solches gilt es die Gründe und auch den Zeitpunkt von Abbrüchen detailliert zu dokumentieren, um nuancierte Auswertungen zu ermöglichen.

Reliable Monitoring-Daten sind eine Grundlage, um Gelingen zu beschreiben, d. h. Daten, die den Zeitpunkt von und die Gründe für Statusveränderungen erfassen, ermöglichen es, Verläufe für alle Teilnehmerinnen aufzuzeichnen und zu untersuchen, welche Faktoren zu einem potenziellen Erfolg im Projekt beitragen. Die zentrale Rolle solcher Daten wurde auch für das SoWi-Team zunehmend deutlich, was im engen Austausch mit dem Evaluationsteam zu befriedigenden Lösungen zur Erfassung, Dokumentierung und Auswertung von Daten führte. Über die Zielsetzung von '50 erfolgreichen Teilnehmerinnen' hinausgehend, werden zudem Bemühungen angestellt, die differenzierteren sowie sekundären Zielsetzungen zu erfassen. So erhebt die Evaluation umfassende Daten zu den psychosozialen Belastungen und Ressourcen der Frauen und der Fortschritt der Teilnehmerinnen in Bezug auf die Ziele des Projekts wird nicht binär dokumentiert, sondern ordinal (gar nicht erreicht/teilweise erreicht/erreicht).

Vergleiche anstellen

Die Evaluation beinhaltet qua Design keine Kontrollgruppe, d. h. Vergleiche mit Frauen, die keine oder andere Maßnahmen erhalten, können nicht angestellt werden. Trotzdem ist es möglich Vergleichsgruppen zu identifizieren bzw. statistisch zu kreieren und damit eine Annäherung an die Gelingensfrage zu ermöglichen. Im Folgenden werden mehrere Optionen skizziert:

⁴ Die Evaluation ist für alle Frauen zugänglich, die das Projekt beginnen. Die Rücklaufquote zur Zeit des Zwischenberichts lag bei 70,7-82,0% je nach Definition der Gesamtgröße (d. h. ob mit oder ohne frühe 'Dropouts').

- Verfügbare Daten zur Arbeitsmarktintegration von Frauen mit Migrationshintergrund sowie zu Projekten mit ähnlichen Zielsetzungen wie SoWi können aufgearbeitet werden, um Vergleiche anzustellen. Eine systematische Review zu Arbeitsmarktintegrationsprojekten die Zielgruppe betreffend konnte jedoch nicht gefunden werden, wäre aber hilfreich für diese Zwecke.
- Im Sinne einer Fallkontrollstudie kann der differenzierte Status von Teilnehmerinnen am Ende des Projekts mithilfe von multivariablen Regressionsmodellen (u. a. binäre oder multinomiale logistische Regressionsanalysen) untersucht werden. Offensichtlich sind Vergleiche zwischen Frauen, die die Ziele des Projekts erreicht haben versus denen, die frühzeitig aus dem Projekt ausscheiden bzw. Ziele teilweise erreichen. Systematische Unterschiede zwischen 'Statusgruppen' ermöglichen Auswertungen bezüglich prädiktiven Faktoren und können somit die Frage beantworten, für welche Frauen das Projekt wirksam ist.⁵
- Wie bedeutsam Coaching als Projektelement für das Gelingen im Projekt ist, ist eine Frage von zentralem Interesse für SoWi, wird aber nur schwer zu beantworten sein. In der Wirkungsforschung sind Komponentenstudien von großer Bedeutung, da sie Aufschluss über die Wirkung einzelner Elemente von komplexen Programmen, Interventionen oder Methoden geben. Solcher Art Studien sind gewöhnlich das Resultat von Jahren rigoros durchgeführter Erhebungen. Trotzdem kann es möglich sein, den Grad der Involvierung oder die Qualität des Coaching ausreichend zu beschreiben und als Moderatorvariable in die Auswertung einzubinden, was Aussagen darüber ermöglichen würde, wie viel Variabilität im Gelingen durch Coaching erklärt wird.
- Letztlich können vergleichende Analysen durchgeführt werden, die Frauen in Bezug auf die Zeitdauer bis zum Erfolg oder Abbruch untersuchen (u. a. durch Ereigniszeitanalysen). Auch dieser Ansatz kann Aufschluss über Merkmale von Frauen geben, die eine höhere Wahrscheinlichkeit haben, früher oder später aus dem Projekt auszuschneiden bzw. mehr oder weniger Zeit brauchen, um Zielsetzungen zu erreichen.

Konzepte verschriftlichen

Eine grundsätzliche Voraussetzung für Wirkungsstudien sind theoretisch fundierte und spezifizierte Projektkonzepte. Für Forschungszwecke dienen Leitfäden (Manuale) dazu, ein Handlungskonzept ausreichend zu spezifizieren, d. h. die 'Black Box der Intervention' zu öffnen und zuzusichern, dass ein Projekt adhärent umgesetzt und somit tatsächlich eine Nettowirkung erfasst wird. Leitfäden spielen darüber hinaus eine signifikante Rolle in der Schulung sowie dezidierten Wiederholbarkeit eines Konzepts und können darüber hinaus der Qualitätskontrolle dienen (Addis & Cardemil, 2006, S. 132 f.). Vonseiten der Evaluation wurde die Bedeutsamkeit eines Handlungsleitfadens für das Projekt, insbesondere für das Coaching, frühzeitig betont und

⁵ Das Evaluationsteam ging dieser Frage in einer präliminären Analyse nach und untersuchte Unterschiede zwischen aktiven Teilnehmerinnen und Teilnehmerinnen, die frühzeitig aus dem Projekt ausgeschieden waren, in Bezug auf psychosoziale Belastungen und Ressourcen. Ergebnisse zeigten, dass Teilnehmerinnen mit größeren psychosozialen Ressourcen eine höhere Wahrscheinlichkeit vorwiesen, frühzeitig aus dem Projekt auszuschneiden. Ergebnisse generierten interessante Hypothesen für Folgerhebungen (James et al., 2021).

inzwischen bestehen Bemühungen, eben jenen zu entwickeln, um eine Grundlage für ähnliche Projekte bilden zu können.

Fazit

Am Beispiel eines aktuellen Arbeitsmarktintegrationsprojekts und seiner wissenschaftlichen Begleitstudie konnten Herausforderungen in der Definition von Gelingen vonseiten unterschiedlicher Akteur*innengruppen beschrieben werden. Zielsetzungen entsprechen der jeweiligen Rolle der Akteur*innen und reflektieren unterschiedliche Interessen und Schwerpunktsetzungen. Schwerwiegender für die Beschreibung der Wirksamkeit eines Projekts sind unterschiedliche Auffassungen darüber, ab wann ein Projekt 'tatsächlich' als wirksam beschrieben werden kann. Die Wissenschaft bestimmt Wirkung aufgrund von wissenschaftlichen Gütekriterien und trifft häufig auf eine Praxis, die unvorbereitet für diese Standards ist. Gleichwohl ist die Wissenschaft vielfach unzureichend sensibel für die komplexen internen und externen Prozesse, die die Praxis gestalten und die Implementierung eines Projekts beeinträchtigen. Dieser Beitrag zeigte auf, wie Lösungen im engen Austausch zwischen Praxis und Wissenschaft entwickelt werden können, die zu größerer forschungsmethodischer Sicherheit führen und zugleich den Realitäten der Praxis gerecht werden.

Wirkmodelle in der Sozialen Arbeit müssen zielgruppenspezifischen Rahmenbedingungen und rechtlichen und gesellschaftspolitischen Kontextfaktoren Beachtung schenken (Bartsch, Beywl & Niestroj, 2016, S.89 f.), was bedeutet, dass das Spektrum der methodischen Möglichkeiten, mit dem Sozialarbeitswissenschaftler*innen sich der Wirkungsfrage annähern, ausgeschöpft werden muss, dass der Erkenntnisprozess als iterativ verstanden werden sollte und dass die Frage für wen und unter welchen Bedingungen ein Projekt wirksam ist, von größerer Bedeutung sein mag als die Determinierung der Nettowirkung einer Intervention.

In der Gesamtschau der Ziele kann die Zusammenarbeit dazu führen, dass auf Grundlage einer empirisch-wissenschaftlichen Begleitung sowie handlungspraktischer Erfahrungen der Praxiskräfte ein (Modell-)Projekt von der Einmaligkeit in ein Regelangebot überführt werden kann und sich daraus eine Kontinuität im Spektrum sozialarbeiterischer Unterstützungsangebote etabliert.

Prof. Dr. Sigrid James ist seit 2016 Professorin am Fachgebiet Theorien und Methoden der Sozialpädagogik der Universität Kassel. Von 2004 bis 2016 hatte sie einen Lehrstuhl an der Loma Linda University in Kalifornien, USA. Ihr Forschungsinteresse gilt Handlungskonzepten und Methoden in der Sozialen Arbeit, insbesondere der evidenzbasierten Praxis. Neben Evaluationen im Bereich der Migrationsforschung widmet sich ihre Forschung auf internationaler Ebene den stationären Hilfen zur Erziehung.

Franziska Seidel ist wissenschaftliche Bedienstete und Doktorandin am Fachgebiet Theorien und Methoden der Sozialpädagogik der Universität Kassel und Koordinatorin der Evaluationsstudie. Forschungsinteressen sind neben Migrationsforschung, Transnationale Familienbeziehungen, Internationale Soziale Arbeit und Kinder- und Jugendhilfe.

Julian Trostmann ist wissenschaftlich Bediensteter am Fachgebiet Erziehungswissenschaften mit dem Schwerpunkt Sozialpädagogik und außerschulische Bildung der Universität Kassel. Forschungsinteressen sind neben qualitativen

und quantitativen Sozialforschungsmethoden, dabei insbesondere in Mixed-Methods-Designs, kulturelle Bildung insbesondere in peripheren Räumen.

Terhas Andezion ist studierte Wirtschaftswissenschaftlerin und als Angestellte der Stadt Kassel leitet sie seit März 2019 das Projekt "Sozialwirtschaft integriert".

Literatur

- Addis, M. E. & Cardemil, E. V. (2006). Does manualization improve therapy outcomes? In J. C. Norcross, I. E. Beutler, R. F. Levant (Hrsg.), *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions* (S. 131-160). Washington, DC., USA: APA.
- Aichberger, M.C., Bromand, Z., Rapp, M.A., Yesil, R., Montesinos, A.H., Temur-Erman, S., Heinz, A., & Schouler-Ocak, M. (2015). Perceived ethnic discrimination, acculturation, and psychological distress in women of Turkish origin in Germany. *Social Psychiatry and Psychiatric Epidemiology*, 50(11), 1691–1700. <https://doi.org/10.1007/s00127-015-1105-3>
- Antidiskriminierungsstelle des Bundes (2020). *Rassistische Diskriminierung auf dem Wohnungsmarkt. Ergebnisse einer repräsentativen Umfrage*. Berlin.
- Aumüller, J. (2016). *Arbeitsmarktintegration von Flüchtlingen: Bestehende Praxisansätze und weiterführende Empfehlungen*. Gütersloh: Bertelsmann Stiftung.
- Bartsch, S., Beywl, W., & Niestroj, M. (2016). Der Programmbaum als Evaluationsinstrument. In S. Giel, K. Klockgether, & S. Mäder (Hrsg.), *Evaluationspraxis. Professionalisierung – Ansätze – Methoden* (2. Aufl., S. 89-113). Münster: Waxmann.
- Bleck, C. (2016). Qualität, Wirkung oder Nutzen: Zentrale Zugänge zu Resultaten Sozialer Arbeit in professionsbezogener Reflexion. In S. Borrmann & B. Thiessen (Hrsg.), *Wirkungen Sozialer Arbeit: Potentiale und Grenzen der Evidenzbasierung für Profession und Disziplin* (S. 107-124). Opladen, Berlin & Toronto: Verlag Barbara Budrich.
- Bundesgesundheitsministerium (2021). *Beschäftigte in der Pflege*. Zugriff am 25.03.2021. Verfügbar unter: <https://www.bundesgesundheitsministerium.de/themen/pflege/pflegekraefte/beschaeftigte.html#c3332>
- Bundesministerium für Familie, Senioren, Frauen und Jugend (2018). *So gelingt der Berufseinstieg von geflüchteten Müttern. Erkenntnisse aus dem ESF-Bundesprogramm „Stark im Beruf – Mütter mit Migrationshintergrund steigen ein“*. Berlin.
- Bundesministerium für Familie, Senioren, Frauen und Jugend (2019). *POINT – Potentiale integrieren. Ergebnisse eines gleichstellungspolitischen Modellprojekts zur Integration allein geflüchteter Frauen in Arbeit und Ausbildung*. Berlin.
- Die Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration (2019). *Deutschland kann Integration: Potenziale fördern, Integration fordern, Zusammenhalt stärken. 12. Bericht der Beauftragten der Bundesregierung für Migration, Flüchtlinge und Integration*. Frankfurt.
- Die Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration (2021). *Einblicke in die Projektarbeit*. Zugriff am 25.03.2021. Verfügbar unter: <https://www.integrationsbeauftragte.de/ib-de/themen/projekte-und-forschung/projekte>

- Die Bundesregierung (2007). Der Nationale Integrationsplan. Neue Wege – Neue Chancen. Berlin: Presse- und Informationsamt der Bundesregierung.
- Die Bundesregierung (2020). Fachkräfteeinwanderungsgesetz. Mehr Fachkräfte für Deutschland. Zugriff am 05.01.2021. Verfügbar unter <https://www.bundesregierung.de/bregde/aktuelles/fachkraefteeinwanderungsgesetz-1563122>
- Dobischat, R. & Schäfer, A. (2020). Projektevaluation „SmartSt@rt“ (3. Zwischenbericht) für den Zeitraum vom 01.07.2019 bis zum 31.01.2020. https://www.uni-due.de/imperia/md/images/biwi/einrichtungen/bawb/3_evaluationbericht_ude_31.01.2020.pdf
- Esser, H. (2006). Migration, Sprache und Integration (AKI-Forschungsbilanz, 4). Berlin: Wissenschaftszentrum Berlin für Sozialforschung gGmbH FSP Zivilgesellschaft, Konflikte und Demokratie Arbeitsstelle Interkulturelle Konflikte und gesellschaftliche Integration.
- Fachkommission Integrationsfähigkeit (2021). Gemeinsam die Einwanderungsgesellschaft gestalten. Bericht der Fachkommission der Bundesregierung zu den Rahmenbedingungen der Integrationsfähigkeit, Berlin: Bundeskanzleramt.
- Farrokhzad, S. (2006). Exotin, Unterdrückte und Fundamentalistin. In C. Butterwegge & G. Hentges (Hrsg.), Massenmedien, Migration und Integration: Herausforderungen für Journalismus und politische Bildung (Bd. 17, 2. Aufl., S. 55–86). Wiesbaden: Springer.
- Farrokhzad, S. (2018). Qualifikation und Teilhabe geflüchteter Frauen und Männer am Arbeitsmarkt. In R. Ceylan, M. Ottersbach & P. Wiedemann (Hrsg.), Neue Mobilitäts- und Migrationsprozesse und sozialräumliche Segregation (S.157–185). Wiesbaden: Springer.
- Ferenschild, S. (2019). Vergeudete Talente. Migrantinnen in Deutschland und berufliche Integration. Bonn: Südwind Institut für Ökonomie und Ökumene.
- Gollwitzer, M. & Jäger, R. S. (2014). Evaluation kompakt (2. Aufl.). Weinheim Basel: Beltz.
- Grünheid, R. & Sulak, H. (2016). Bevölkerungsentwicklung Daten, Fakten, Trends zum demografischen Wandel. Wiesbaden: Bundesinstitut für Bevölkerungsforschung.
- Habermann, M. & Stagge, M. (2015). Menschen mit Migrationshintergrund in der professionellen Pflege. In P. Zängl (Hrsg.), Zukunft der Pflege: 20 Jahre Norddeutsches Zentrum zur Weiterentwicklung der Pflege (Bd. 4, S. 159–175). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-08137-9_14
- Hessisches Ministerium für Soziales und Integration (2018). Hessischer Integrationsplan. Für eine Kultur des Miteinanders in Respekt und gegenseitige Anerkennung. Zugriff am 25.03.2021. Verfügbar unter https://soziales.hessen.de/sites/default/files/media/hsm/hessischer_integrationsplan.pdf
- Hessisches Ministerium für Soziales und Integration (n.d.). Sozialwirtschaft integriert. Ausbildung ist der beste Anfang: 10 Projekte, die Chancen schaffen (Arbeitswelt Hessen, Hrsg.). Wiesbaden.
- Holosko, M. J. (2010). What types of designs are we using in social work research and evaluation? *Research on Social Work Practice*, 20(6), 665–673.
- Irfaeya, M., Maxwell, A.E., & Krämer, A. (2008). Assessing psychological stress among Arab migrant women in the city of Cologne/Germany using the Community Oriented Primary Care (COPC) approach. *Journal of Immigrant and Minority Health*, 10(4), 337–344. <https://doi.org/10.1007/s10903-007-9091-5>

- James, S. (2019). Wirkungsforschung in der Kinder- und Jugendhilfe – Internationale Perspektiven. In M.-C. Begemann, C. Bleck, R. Liebig (Hrsg.), *Wirkungsforschung zur Kinder- und Jugendhilfe* (S. 45-67). Weinheim: Beltz Juventa.
- James, S., Seidel, F.A., Kilian, J., & Trostmann, J. (2020). Labor market integration of young adult refugees in Germany – triangulating perspectives toward program development. *Research on Social Work Practice*, 30(5), 553-563.
- James, S., Seidel, F.A., & Trostmann, J. (2021). Psychosoziale Belastungen und Ressourcen von Frauen mit Migrationshintergrund in einem städtischen Arbeitsmarktintegrationsprojekt. In A. Holleder (Hrsg.), *Gesundheitsförderung bei Arbeitslosen* (S. 365-380). Frankfurt a.M.: Fachhochschulverlag.
- Kluge, U., Rapp, M.A., Mehran, N., Jumaa, J., & Aichberger, M.C. (2019). Armut, Migration und psychische Gesundheit. *Der Nervenarzt*, 90(11), 1103-1108. <https://doi.org/10.1007/s00115-019-00790-2>
- Kränzl-Nagl, R., Lehner, M., & Prinz, T. (2019). *Sozialökonomische Wirkungsevaluation in der Sozialwirtschaft: Grundlagen und Praxisbeispiele*. Regensburg: Walhalla.
- Lukas, W. (2011). *Migranten im Niedriglohnsektor unter besonderer Berücksichtigung der Geduldeten und Bleibeberechtigten*–Working Paper 39. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Merchel, J. (2019). *Evaluation in der Sozialen Arbeit* (3. Aufl.). Stuttgart: utb.
- Nüsken, D. (2015). Wie breit ist das Feld? Stand der Wirkungsforschung und ein kleiner Überblick, was alles Evaluation und Wirkungsforschung ist. In Arbeitsgruppe Fachtagungen Jugendhilfe im Deutschen Institut für Urbanistik (2015). *Aktuelle Beiträge zur Kinder- und Jugendhilfe* 100. Berlin: Deutsches Institut für Urbanistik gGmbH.
- Oestreicher, E. (2014). *Wissenstransfer in Professionen. Grundlagen, Bedingungen und Optionen*. Opladen: Budrich UniPress Ltd.
- Ottmann, S. & König, J. (2018). 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. Forschung, Entwicklung, Transfer – Nürnberger Hochschulschriften*, Nr. 29. Nürnberg: Evangelische Hochschule Nürnberg. doi: 10.17883/fet-schriften029.
- Sachverständigenrat deutscher Stiftungen für Integration und Migration (2019). *Bewegte Zeiten: Rückblick auf die Integrations- und Migrationspolitik der letzten Jahre. Jahresgutachten 2019*. Berlin: Sachverständigenrat deutscher Stiftungen für Integration und Migration (SVR) GmbH.
- Salikutluk, Z., Giesecke, J., & Kroh, M. (2020). The situation of female immigrants on the German labour market: A multi-perspective approach. Berlin: SOEP Papers on Multidisciplinary Panel Data Research.
- Schäfer, A., & Dobischat, R. (Hrsg.) (2021). *Beruflich-sprachliche Qualifizierung von Geflüchteten als eine Voraussetzung für die Arbeitsmarktintegration. Erkenntnisse, Erfahrungen und Perspektiven aus dem Projekt „Smart St@rt“*. Essen.
- Schneider, A. (2016). Konzepte der Wirkungsmessung und -forschung. *Soziale Arbeit*, 65(6/7), 204-211.

Statistisches Bundesamt (2019). Jede vierte Person in Deutschland hatte 2018 einen Migrationshintergrund. 52% der Menschen mit Migrationshintergrund sind deutsche Staatsangehörige. Zugriff am 05.01.2021. Verfügbar unter https://www.destatis.de/DE/Presse/Pressemitteilungen/2019/08/PD19_314_12511.html

Walter, T., Schütz, H., Bonin, H., Butschek, S., Knerr, P., Schröder, H., & Steinwede, J. (2014). Abschlussbericht Evaluation „Programm zur berufsbezogenen Sprachförderung für Personen mit Migrationshintergrund (ESF-BAMF-Programm)“. Mannheim.

Einleitung

Hochschulen für Angewandte Wissenschaften (HAWs) spielen für den Transfer zwischen Wissenschaft und Praxis, also für die Vermittlung wissenschaftlicher Erkenntnisse und ihrer Bedeutung für den jeweiligen beruflichen Praxisalltag, eine große Rolle. Dabei geht es nicht nur darum, dass entsprechende praxisrelevante Forschung an den HAWs stattfindet – diese ist Teil ihres öffentlichen „Auftrages von Forschung und Lehre“ – sondern vor allem darum, dass den Studierenden die Hintergründe wissenschaftlicher Methodik und die Auseinandersetzung mit wissenschaftlichem Denken und Arbeiten vermittelt werden. Dies hängt eng mit dem Professionalisierungsgedanken der Sozialen Arbeit als eigenständiger (angewandter) Wissenschaft zusammen. Die Notwendigkeit dieser Lehrinhalte erscheint Studierenden (nicht nur an HAWs) hingegen oftmals fraglich: Wozu Wissenschaft, wenn man später doch praktisch arbeiten wolle? Das Studium an einer HAW ist für viele Studierende oft primär der Weg zu einem entsprechenden Berufsabschluss; „Wissenschaft“ und „Wissenschaftlichkeit“ werden dabei als notwendiges Übel auf diesem Weg betrachtet. Diese beiden Komplexe – Wissenschaft auf der einen und Praxis auf der anderen Seite – scheinen im Berufsalltag kaum Berührungspunkte zu bieten. So berichtet Müller (2017): „Gewöhnlich antworten Studienanfängerinnen auf die Frage, weshalb sie sich ausgerechnet das Fach Soziale Arbeit gewählt haben, sie wollen ‚etwas mit Menschen zu tun haben‘“ (ebd., S. 194) – und nicht, dass sie forschen wollen. Diese Einstellung wird in studienbegleitenden Praktika unterstützt, wenn Studierende hören: „In der Theorie (= Wissenschaft) sollte man das so machen, aber hier in der Praxis machen wir das anders.“

Das Thema ist nichts Neues und besteht auch nicht erst, seit 1999 der Bologna-Prozess startete: „Die Klage, das sozialpädagogische beziehungsweise sozialarbeiterische Studium sei praxisfern, ist so alt wie die Ausbildung selbst“ (Müller, 2017, S. 11). Erschwert wird die Situation auch dadurch, dass viele Professor*innenstellen in der Sozialen Arbeit durch Vertreter*innen anderer (universitärer) Fachdisziplinen besetzt sind, die jeweils eigene (wissenschaftliche) Berufsbiografien mitbringen, eigene Fachdisziplinen-orientierte Forschungsmethoden, sowie eine eigene Fachsprache. Wie also kann das Thema „Wissenschaft“ an Hochschulen für Angewandte Wissenschaft besser, anders oder überhaupt vermittelt werden? Wo genau liegen die Probleme? Welche Lösungsideen gibt es? An welchen Punkten kann angesetzt werden, um diesen gefühlten Abstand zwischen „Wissenschaft“ und „Praxis“ zumindest etwas zu verringern?

Im Folgenden sollen drei mögliche Ansätze vorgestellt werden, bevor auf eine Erhebung und Diskussion des Themas an der Ev. Hochschule Rheinland-Westfalen-Lippe (EvH R-W-L) eingegangen wird.

Aus der „Theorie“

Ansatz 1 greift die „typischen“ möglichen beruflichen Haltungen von Sozialarbeitenden auf (vgl. Müller, 2017), die übertragbar sind auf mögliche Haltungen der Studierenden bereits in ihrem Studium. Ansatz 2 beschäftigt sich ebenfalls mit der Frage von Haltungen und Einstellungen, hier zum Thema Lernen und Lernverhalten (vgl. Schrader, 2008). Und Ansatz 3 fokussiert das „Spannungsfeld der Lebens- und Arbeitswelt Hochschule“ (Centeno García & Kenneweg, 2019). Zum ersten Ansatz: Müller (2017) nennt drei „Typen“ eines sozialarbeiterischen Professionsverständnisses, welche hier als Haltungen bezeichnet werden sollen (s. Tab. 1).

Tabelle 1: mögliche Haltungen in der Profession der praktischen Sozialen Arbeit (Müller, 2017, S. 196 ff.)

„geschlossene Haltung“	übernommene Deutungen, routinierte Vorgehensweisen ersetzen eigene Wahrnehmungsfähigkeit; unreflektiertes Verständnis von Professionalität
„autistische Haltung“	die „jeweils existierenden Praktiken eines Berufsfeldes [werden] zum Maß aller Dinge, [...] jede Kritik [wird] als nicht praktikabel, überfordernd, zu zeitaufwändig etc.“ (ebd., S. 199) abgewehrt
„offene Haltung“	„Wissen und professionseigene Arbeitsregeln, die helfen, das je nach Situation und Fall benötigte Handlungswissen selbst zu erschließen“ (ebd., S. 199)

Bei der „geschlossenen Haltung“ beispielsweise wird Professionalität nicht reflektiert; Müller (2017) bezeichnet diese als eine „bestimmte Betrachtungsweise“, welche „kultiviert und mit methodischen Instrumentarien versehen [wird], die zugleich als Ausblendfilter für andere Sichtweisen funktionieren“ (ebd., S. 197). Bei der „autistischen Haltung“ andererseits ersetzt die Pragmatik des „gesunden Menschenverstands“ eine theorie-basierte Reflexion und verfestigt damit eine Haltung, „die sich nur noch an dem, was ohnehin läuft, orientiert und [...] gegen jede Kritik abschirmt“ (ebd., S. 199); die „jeweils existierenden Praktiken eines Berufsfeldes [werden] zum Maß aller Dinge, [...] jede Kritik [wird] als nicht praktikabel, überfordernd, zu zeitaufwändig etc.“ (a.a.O.) abgewehrt. Müller (2017) weiter: „[E]ine solche Professionalität hat keinen Inhalt mehr. Sie verteidigt nur das Recht, keinen Maßstäben genügen zu müssen und keine Rechenschaft für ihr Handeln schuldig zu sein“ (ebd., S. 199). Die „offene Haltung“ hingegen bemüht sich um die Berücksichtigung spezifischer Bedingungen eines Handlungsfeldes, um Zieloffenheit und um die Fähigkeit zum Perspektivwechsel zwischen „Wissen und professionseigene Arbeitsregeln, die helfen, das je nach Situation und Fall benötigte Handlungswissen selbst zu erschließen“ (ebd., S. 199).

Bezogen auf das Studium kann mit diesen drei Haltungen der Entwicklungsprozess der Studierenden verdeutlicht werden. So wird gerade zu Studienbeginn eher ein „geschlossener“ Fokus gefunden, wenn zum Beispiel die Inhalte der verschiedenen human- und sozialwissenschaftlichen Grundlagendisziplinen strikt nach Fachdisziplin getrennt gelernt werden: In der Pädagogik ist dies Perspektive A, in der Psychologie Perspektive B, in der Soziologie Perspektive C, und alle drei stehen losgelöst voneinander. Die „richtige“ Perspektive bestimmt sich für den*die einzelne*n Studierende*n dann aufgrund der Präferenz eines*einer Lehrenden, der

vermuteten Sinnhaftigkeit für die berufliche Praxis, oder anderen Motiven. Zur „autistischen Haltung“ kann es dann kommen, wenn Studierende in ihren Praktikumsphasen die bereits benannte (vermeintliche) Theorie-Praxis-Diskrepanz erleben. Die „offene Haltung“ finden Studierende dann gegebenenfalls, wenn sie sich im Rahmen ihrer Abschlussarbeit mit einer anwendungsbezogenen Fragestellung, ihrer theoretischen Grundlage und dem entsprechenden Rückbezug auf die Praxis auseinandersetzen. Im Zusammenhang mit der hier behandelten Frage nach der Wissenschaftlichkeitsvermittlung wäre es wichtig, diese Haltungen und ihre Entwicklungen in den Lehrveranstaltungen vorzustellen, zu diskutieren und zu reflektieren. Diese sich im Studium entwickelnden Haltungen hängen von bestimmten Einstellungen bezüglich des eigenen Lernverhaltens ab. Schrader (2008, S. 110) nennt diese „Typologien von Lernformen“; im Kontext der Ausführungen des vorliegenden Papers wird der Begriff der „Haltungen“ verwendet (s. Tab. 2). Daneben wird davon ausgegangen, dass diese Haltungen sich nicht per se und in allen Seminaren und Modulen gleich zeigen, sondern dass auch hier ein Situationsbezug besteht.

Tabelle 2: mögliche Lernhaltungen im Studium (Schrader, 2008, S. 110 ff.)

„theoretische“ Lernhaltung	Freude am Lernen; konkrete Vorstellungen von dem, was er*sie lernen will; nicht nur an praktischer Anwendung, sondern auch an theoretischen Grundlagen interessiert
„anwendungsorientierte“ Lernhaltung	„Was kann ich mit neuen Inhalten anfangen?“; Theorien und reines Faktenwissen genügen nicht
Lernhaltung „Musterschüler“	lernt für gute Noten, für Zeugnisse und Zertifikate; Einprägen möglichst vieler Fakten, um diese zu einem späteren Termin möglichst gut wiedergeben zu können; Schwierigkeiten, wenn es keine eindeutige Lösung gibt
„gleichgültige“ Lernhaltung	ökonomisches Lernen: nicht mehr, als unbedingt erforderlich, um nicht auf- oder durchzufallen
„unsichere“ Lernhaltung	Lernen begleitet von Unsicherheit und Angst; Erwartung zahlreicher Schwierigkeiten beim Lernen

Sicherlich erhoffen sich Lehrende (je nach eigener wissenschaftlicher Studierbiografie) vor allem Studierende mit einer „theoretischen“ oder „anwendungsorientierten“ Lernhaltung in möglichst allen Lehrveranstaltungen. Allerdings sollte die Quote an den anderen studentischen Lernhaltungen nicht unterschätzt werden – abhängig von den Lernerfahrungen, die zum Beispiel in der Schule in verschiedenen Fächern gemacht wurden, wird auch die Lernhaltung in den Seminaren und Vorlesungen an der Hochschule beeinflusst. „Lernen“ ist in seiner Gesamtheit ein individueller Prozess (vgl. u.a. Wahl, 2020, S. 101, Abb. 17), beeinflusst auch von individuellen, situationsbezogenen Lernhaltungen.

Die beiden bisher genannten Perspektiven – die der beruflich-studentischen Haltung zur Profession einerseits und der studentischen Lernhaltung an sich andererseits – finden eine Problemverdoppelung, wenn der Kontext Hochschule genauer betrachtet wird (Huber, 1983, S. 128, zit. n. Centeno García & Kenneweg, 2019, S. 900): Es gibt nicht nur die eine Wissenschaftshaltung und die eine Praxishaltung, nicht nur die eine Person mit ihren individuellen Lehr-Lernhaltungen, sondern es stellt sich die Frage nach der jeweiligen Positionierung (s. Abb. 1).

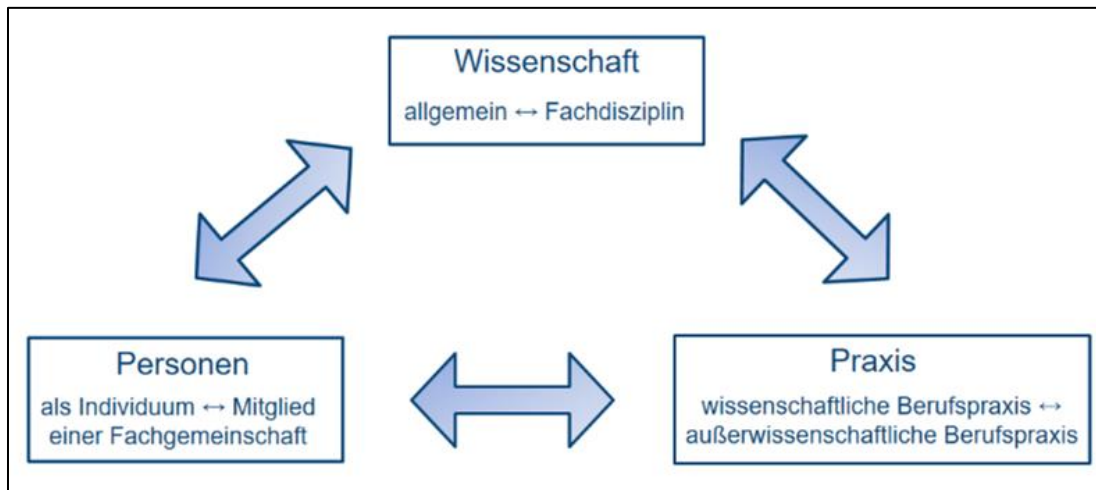


Abbildung1: Spannungsfeld der Lebens- und Arbeitswelt Hochschule (s. Centeno García & Kenneweg, 2019, S. 900, Abb. 2).

So setzt jede Wissenschaftsdisziplin neben einer generellen Verständigung über allgemeingültige wissenschaftliche Standards ihre eigenen Schwerpunkte innerhalb der fachdisziplinarischen Methodik. „Praxis“ gibt es sowohl im Kontext von Wissenschaft als auch im außerwissenschaftlichen Kontext. Und jede Lehrperson ist nicht nur ein*e Vertreter*in der eigenen Fachdisziplin mit ihren Standards, sondern auch noch ein Individuum mit entsprechender Berufsbiografie, Lehr-Lernhaltung und Sichtweise auf die Bedeutsamkeit der eigenen Disziplin für das entsprechende Studienfach. Wie bereits einleitend erwähnt, bildet die Zusammenarbeit mit professoralen Vertreter*innen anderer (universitärer) Fachdisziplinen eine besondere Herausforderung hinsichtlich der zum Teil fachfremden Vermittlung des Themas „Wissenschaft und Praxisbezug in der Sozialen Arbeit“.

Um sich **der** Wissenschaft zu nähern, stellt sich also zuerst die Frage nach dem kleinsten gemeinsamen Nenner in den vielen verschiedenen (sozialwissenschaftlichen) Fachdisziplinen hinsichtlich einer Definition von „Wissenschaft“ bzw. „Wissenschaftlichkeit“. In entsprechender Grundlagenliteratur zum wissenschaftlichen Arbeiten findet sich hier oftmals eine Eingrenzung vor allem auf eine empirische Perspektive. Offener hingegen definierten dies Brauner und Vollmer (2008):

„Eine Arbeit kann als wissenschaftlich gelten, wenn sie

- einem Fachgebiet zugeordnet ist,
- die Terminologie des Fachgebietes korrekt verwendet und
- weitere Konventionen und Richtlinien erfüllt“ (ebd., S. 18).

Bohl (2008, S. 11) zählt als Kriterien wissenschaftlichen Arbeitens auf: „eigenständige Gedankenarbeit; systematisches und methodisch kontrolliertes Vorgehen; Allgemeingültigkeit; Fundierung der Aussagen; Schreibstil; Begriffsklarheit; formale und technische Aspekte; Redlichkeit.“ Bastian und Groß-Mlynek (2018) definieren als Ziel wissenschaftlichen Arbeitens die Fähigkeit, „unbekannte Probleme selbstständig zu lösen“ (ebd., S. 25), wozu gehört: „Probleme wahrnehmen; neugierig sein, Fragen stellen und genau beobachten; Systematisieren und

Differenzieren; andere Meinungen einbeziehen; kritisch lesen und verstehen; Wissen einordnen und sichern" (a.a.O.). Als Minimalkonsens lässt sich letztendlich die Transparenz und Nachvollziehbarkeit der jeweiligen Argumentationsstrukturen festhalten: Wie kommt ein bestimmtes Ergebnis zustande?

Was bedeutet dies nun für die „Wissenschaftliche Soziale Arbeit“? Mennemann und Dummann (2020) erläutern dies folgendermaßen:

„Soziale Arbeit als Profession und als Disziplin [...] ermöglicht, reflektiert und beforscht die Gestaltung sozialer Begegnungsräume, die am Bedarf der Adressat*innen lebensnah ausgerichtet werden, um die Bewusstseins- und Handlungspotentiale so zu stärken oder die Umweltbedingungen so zu verändern, dass soziales Miteinander beziehungsweise Teilhabe möglich wird" (ebd., S. 89).

In dieser Erläuterung wird deutlich, dass nicht nur ein „Ermöglichen“ und ein „Reflektieren“, sondern eindeutig auch ein „Beforschen“ Teil der Disziplin der Sozialen Arbeit ist. Umso mehr stellt sich die Frage, wie es zur wahrgenommenen Diskrepanz zwischen „Theorie und Forschung“ auf der einen und „Praxis“ auf der anderen Seite kommt: Woran misst sich eine in die Praxis umsetzbare Theorie? Und mit welchen Methoden kann man dies messen? Welche wissenschaftlichen Forschungsmethoden sind hier anwendbar? Bevor hier nach Antworten gesucht wird, muss zuerst noch eine grundsätzlichere Frage gestellt werden: Sind „Wissenschaft“ und „Forschung“ also synonym zu verwenden? Welche Vorstellung von „Wissenschaft“ und „wissenschaftlicher Arbeit“ haben wir als Lehrende eigentlich, und welche unsere Studierenden?

Die Untersuchung

Zur Annäherung an eine Antwort auf diese Grundsatzfrage wurde ein standardisierter Kurzfragebogen entwickelt (s. Anhang), der ab dem Wintersemester 2019/20 an Studierende verschiedener HAWs mit unterschiedlichen Fachrichtungen und Studiengängen verteilt wurde und sie zu ihrem eigenen Wissenschafts-Verständnis und zum wahrgenommenen Zusammenhang mit dem eigenen Studienfach befragte. Im Sommersemester 2020 wurde die Frage nach der Definition von „Wissenschaft“ auch an verschiedene Lehrende verschiedener Fachdisziplinen gestellt. Die Ergebnisse dieser Erhebungen wurden im Wintersemester 2020/21 mit Studierenden der EvH R-W-L diskutiert und gemeinsam nach Lösungsansätzen gesucht, um „Wissenschaft“ für Studierende allgemein und im Besonderen greifbarer zu machen.

Insgesamt haben 170 Studierende an der Kurzfragebogen-Befragung teilgenommen: 62 Studierende kamen aus dem BA-Studiengang der Sozialen Arbeit (= 36,5%), 40 Personen studierten einen BA-Studiengang aus dem Gesundheitsbereich (= 23,5%), 35 einen BA Softwaretechnik (= 20,6%), 27 einen BA Pflegewissenschaft (= 15,9%), weitere sechs Studierende verteilten sich auf die BA-Studiengänge Heilpädagogik/ Inklusiv Pädagogik, Gemeindepädagogik/ Diakonie und Elementarpädagogik.

Ein erster Überblick über die Verbindung des eigenen Studiengangs mit „Wissenschaft“ deutet zumindest auf unterschiedliche Wahrnehmungen im Zusammenhang mit dem jeweiligen Studienfach hin – die Verzahnung von „Wissenschaft“ mit Informatik scheint eine andere zu sein als mit Sozialer Arbeit oder Pflegewissenschaft (s. Abb. 2).

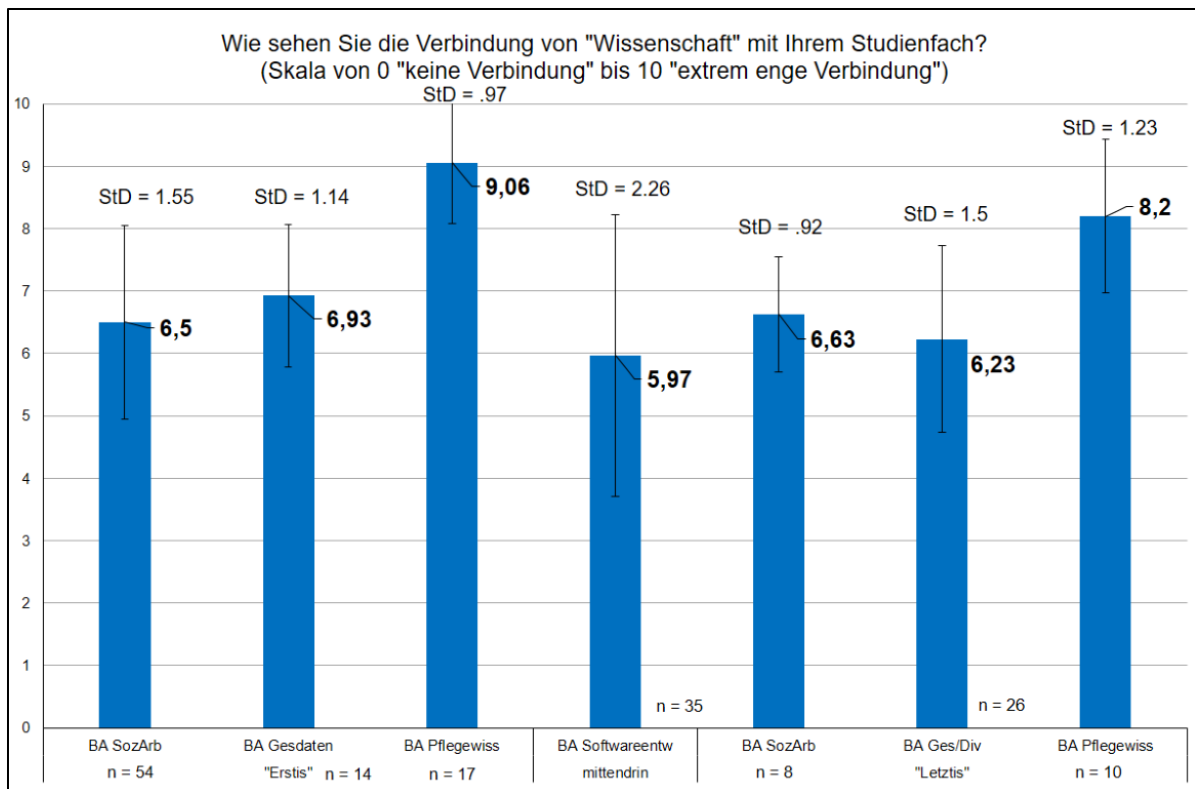


Abbildung 2: Auswertung der zweiten Frage aus dem Kurzfragebogen, aufgeteilt in Studiengänge und Semesterhöhe. Angeben sind jeweils die arithmetischen Mittel (fett) und die Standardabweichungen.

Die Argumente, die die Studierenden als Begründung für die jeweilige Skalierung angaben (Frage 3 des Kurzfragebogens), konnten vier inhaltlichen Hauptkategorien zugeordnet werden:

1. Wissenschaft als Teil des Studiums – in diese Kategorie wurden Antworten einsortiert, die auf ein entsprechendes Modul zum wissenschaftlichen Arbeiten oder auf den Namen des Studiengangs hinwiesen, in dem das Wort „Wissenschaft“ auftauche; weil man wissenschaftliche Texte lesen müsse; etc.
2. Wissenschaft als Basis ging dagegen auch inhaltlichen Bezügen nach: Man beschäftige sich schließlich mit Forschung, sammle Erkenntnisse und stelle Thesen auf, die untersucht werden müssten.
3. Auf praktische Umsetzung verwiesen Äußerungen der dritten Kategorie: Man wende bestimmte „wissenschaftliche“ Methoden an, habe interdisziplinären Austausch, es ginge um aktuelle Studien und Ähnliches.
4. Kritisch betrachtet wurde die Verbindung vor allem von Studierenden aus höheren Semestern – z. B. indem festgestellt wurde, dass doch weniger „Wissenschaftlichkeit“ vermittelt worden sei als erwartet. Speziell von Studierenden der Sozialen Arbeit wurde hervorgehoben, dass die Verbindung nicht so eng sei, weil man in der (praktischen) Sozialen Arbeit ja vor allem Empathie und nicht „Wissenschaft“ benötige.

Was ist also „Wissenschaft“ für Studierende? In Abbildung 3 ist eine einfache Auswertung der entsprechenden offenen Antworten der Studierenden dargestellt: In einer so genannten

Wordcloud werden die Begriffe nach der Häufigkeit ihrer Nennungen größenmäßig präsentiert; das Ergebnis: „Wissenschaft“ für Studierende ist gleich „Forschung“.



Abbildung 3: „Was ist für Sie Wissenschaft?“ – Word Cloud aus allen Studierenden-Nennungen (n = 170) (<https://tagcrowd.com/>; 09.12.2020). Exkludierte Wörter: Bereichen bestimmten etc für können müssen stellen über Wissenschaft wissenschaftliche wissenschaftliches; gezeigt werden die Top 50 von insgesamt 799 möglichen Wörtern mit mind. 4 Nennungen.

Eine tiefere inhaltliche Analyse, in der die genannten Aussagen zu Kategorien zusammengefasst wurden, macht diese Gleichsetzung noch deutlicher¹:

In die **Kategorie „Forschung“** wurden Äußerungen zusammengefasst, die *Forschung, Experimente, Hypothesen, Studien, Forschungsdesign, „hinterfragen“, Auswertung/ Ergebnisse, wissenschaftliches Arbeiten, wissenschaftliche Diskussion, recherchieren, Statistik* fokussierten. Allein die direkte Erwähnung von „Forschung“ tauchte in den Antworten von 97 Fragebögen auf. Aussagen, die der Kategorie „empirisch“ zugeordnet wurden (Themen hier: *Prüfbarkeit, Theorien, Methoden, Wahrheit, belegbar, objektiv, Evidenz*) fanden sich in den Antworten von 81 Fragebögen.

Andere Nennungen wurden zu weiteren Kategorien zusammengefasst (***Bildung; Entdeckung; Alltag; Lösungen schaffen; Vielfalt***), allerdings konnten hierzu weitaus weniger Fragebögen zugeordnet werden.

Was bedeutet „Wissenschaft“ für Lehrende? Auf die entsprechende Frage antworteten insgesamt zehn Lehrende unterschiedlicher Fachdisziplinen. Auch hier geht es also allenfalls um eine Eindrucksbildung, um die Darstellung von Tendenzen und Entwickeln von Hypothesen und nicht um eine statistisch belastbare Aussage. Dennoch zeigt sich ein großer Unterschied hinsichtlich

¹ Aussagen beziehen sich zum Teil auf mehrere Kategorien, daher insgesamt mehr als 170 Fragebögen.

der Lehrenden-Definition von der Studierenden-Definition (s. Abb. 4): „Forschung“ ist nur ein Thema unter vielen anderen.



Abbildung 4: „Was ist für Sie Wissenschaft?“ – Word Cloud aus Lehrenden-Nennungen verschiedener Fachdisziplinen und HAWs (insgesamt 10 Teilnehmende) (<https://tagcrowd.com/>; 09.12.2020). Exkludierte Wörter: *Art Fall Gegenstandsbereich Wissenschaft bzw eher eigentlich für jeweiligen können lassen mal ne über überhaupt verschiedenen wissenschaftliche*; gezeigt werden die Top 50 von insgesamt 442 möglichen Wörtern mit mind. 2 Nennungen.

Die inhaltliche Analyse der Lehrende-Aussagen macht deutlich²: Diese sehen hinter dem Begriff „Wissenschaft“ ein systematisches Vorgehen (5 Befragte), ein Hinterfragen (5 Befragte), ein Anwenden bestimmter Methoden und Techniken (5 Befragte). Es gehe um die Generierung von Wissen (3 Befragte) und bedürfe ausreichendem Durchhaltevermögen (3 Befragte). Daneben wurden von Einzelpersonen noch insgesamt elf weitere Themen genannt, die verdeutlichen: Auf der einen Seite stehen die Lehrenden mit einem sehr breiten, offenen Verständnis von Wissenschaft – und auf der anderen Seite eine Vielzahl von Studierenden mit einem sehr eingegrenzten, geschlossenen Verständnis. Könnte dies Teil des Problems bezüglich der wahrgenommenen Trennung zwischen Wissenschaft und Praxis sein? Gibt es implizite und nicht kommunizierte Unterschiede in Definition und Verständnis von „Wissenschaft“ auf Seiten der Lehrenden und auf Seiten der Studierenden, deren explizite Kommunikation und Diskussion einen Schlüssel zur Lösung bieten könnte? Geht es also eigentlich „nur“ um ein Kommunikationsproblem?

Diese Gedanken wurden zusammen mit den oben vorgestellten Befragungs-Ergebnissen im Wintersemester 2020 mit Studierenden der EvH R-W-L diskutiert. An der ersten von zwei virtuell stattgefundenen Diskussionsrunden nahmen im Dezember 2020 insgesamt sieben Studierende teil³ (fünf Personen aus dem Studiengang BA Soziale Arbeit, eine Person aus dem Studiengang BA Heilpädagogik/ Inklusive Pädagogik und eine Person aus dem Studiengang BA

² Auch hier: Aussagen beziehen sich zum Teil auf mehrere Kategorien, daher insgesamt mehr als 10 Befragte.

³ Das entsprach zu dem Zeitpunkt einer Teilnehmendenquote von 0,27% von allen eingeschriebenen Studierenden an der EvH R-W-L. Die Teilnehmenden konnten mithilfe einer Informations-Email gewonnen werden, welche etwa einen Monat vor der ersten Diskussionsrunde an *alle Studierenden* der Hochschule versendet wurde.

Elementarpädagogik). Eine Person war zu diesem Zeitpunkt im 2. Semester, vier im 5. Semester und zwei im 7. Semester. An der zweiten Diskussionsrunde im Februar 2021 nahmen noch fünf teilnehmende Studierende teil (davon kamen drei Personen aus dem Studiengang BA Soziale Arbeit und jeweils eine Person aus dem Studiengang BA HP/IP und aus dem Studiengang BA EP). In diesem Austauschprozess betonten die Studierenden noch einmal die individuellen Biografien und auch Lerngeschichten der Studierenden an Hochschulen insgesamt und an HAWs vielleicht noch mehr als an Universitäten. Es sei ein (Zitat) „großer Schritt von einfacherer Reproduktion in der Schule (Referate) zu komplexen, belastbaren Zusammenhängen im Studium“ und man bat darum, (Zitat) „noch mehr an die Hand genommen“ zu werden. Für die Studierenden bedeute dies im Zusammenhang mit der Lehre zur Einführung ins wissenschaftliche Arbeiten folgendes: Es solle eine Vereinheitlichung der Inhalte in den Propädeutik-Seminaren stattfinden, so dass die Qualität der Seminare für alle Studierenden möglichst gleich sei. Daneben wäre es hilfreich, wenn Lehrende eine ausführlichere Definition kommunizieren, dass (Zitat) „Wissenschaft nicht nur ‚Forschung‘ ist, sondern ‚empirische, evidenzbasierte Auseinandersetzung““. Diese findet dementsprechend nicht nur losgelöst in einem Propädeutik-Modul zu Beginn des Studiums statt, sondern sollte, so die Forderung der Studierenden, in allen Modulen und Semestern aufgegriffen werden. So könne gerade eine engere Verbindung von „Wissenschaft“ mit den Praxissemestern für eine Verdeutlichung des Zusammenhangs zwischen „Theorie“ und „Praxis“ sorgen. Letztendlich gehe es um ein *Empowerment* der Studierenden, also einer Art von „Hilfe zur Selbsthilfe“, dem Standard praktischer Sozialer Arbeit (s. Schilling, 2020, S. 71; ausführlicher: Lambers, 2020, S. 386ff.): Die Studierenden bitten darum, ihnen (Zitat) „Mut zu machen“ und die (Zitat) „studentische Verunsicherung auszuhalten“. Sie benötigen Zeit zur Reflexion gerade wissenschaftlicher Themen, und sie heben die Problematik der Prüfungen hervor: Ein (Zitat) „freudvolles Auseinandersetzen mit Wissen ist nicht möglich, da immer etwas abzuliefern ist“. Von den Lehrenden erwarten die Studierenden eine wertschätzende und motivierende Grundhaltung; der Ansatz von Müller (2017, S. 12⁴) und von Wahl (2020, S. 19⁵) steht für sie als Wahrnehmung Studierender als kompetenter Personen – (Zitat): „da bleibt man lieber dran in Veranstaltungen“. Lehrende seien daher eben nicht nur für die Fachinhalte verantwortlich, sondern auch für die dazugehörige didaktische Methodik.

Zur Umsetzung in die Lehr-„Praxis“

Einige der oben genannten Probleme und Schwierigkeiten (wie zum Beispiel „Auflösung“ der Fachgruppen und interdisziplinärer Austausch innerhalb der verschiedenen Module; Arbeiten an Themen und nicht mehr Fokussierung auf Einzelprofessionen) sollten eigentlich bereits durch den Bologna-Prozess umgangen und gelöst werden. Dieses kann einigen Hochschulen besser gelungen sein als anderen. Die Gründe hierfür sind vielfältig und nicht Thema dieser Untersuchung. Stattdessen wechsle ich in eine persönliche Betrachtung: Was bedeuten die Ergebnisse der Recherchen, Befragungen und Diskussionen mit Studierenden für *mich* als Lehrende?

⁴ „...[...], dass Hochschullehrer schlecht beraten sind, wenn sie Studierende als ahnungslose und erfahrungslose Wesen behandeln, statt sie als werdende Fachleute mit eigenem Kopf ernst zu nehmen“ (Müller, 2017, S. 12).

⁵ „Was wir wissen und was wir können ist nicht nur Resultat von Lernprozessen. Es ist vor allem der wichtigste Ausgangspunkt für den Erwerb weiteren Wissens und weiterer Kompetenzen!“ (ebd., S.19).

Die Basis ist: Studium ist Erwachsenenbildung. Studierende bringen (Lern- und Lebens-)Erfahrungen mit, die sie im hochschulischen Kontext für sich und andere gewinnbringend einbringen können. Die Verantwortung für das eigene Lernen wird ihnen also nicht abgesprochen. Gleichzeitig: Lehrende sollten sich immer wieder verdeutlichen, dass auch erwachsene Studierende nicht wissen, was sie nicht wissen, wenn ihnen relevantes Vorwissen fehlt; gerade alltägliche Begrifflichkeiten, die in der Sprache einer Fachdisziplin unterschiedliche Bedeutungen haben (können), gaukeln möglicherweise ein Verstehen vor, welches doch nicht existiert.⁶

Zur Illustration: Der Methoden-Begriff taucht im ersten Studiensemester in verschiedenen Kontexten auf – es gibt Methoden in der empirischen Sozialforschung, Methoden in der Psychologie, der Pädagogik, der Soziologie und auch in der Sozialen Arbeit. In dem einen Zusammenhang geht es um Forschungsmethoden qualitativer und/ oder quantitativer Art, in anderen Zusammenhängen um Erziehungsmethoden, Gesprächsführung, Recherchemöglichkeiten, usw. An sich geht es um Erkenntnisgewinn; die Erkenntnis, die Studierende gewinnen, ist vor allem: Es ist verwirrend. Betrachtet man, welche Methoden in der Sozialen Arbeit genannt werden, so finden sich hier beispielhaft bei Mennemann und Dummann (2020, S. 185f.) die Nennungen „Gemeinwesenarbeit – Gruppenarbeit – Einzelfallarbeit“ und ein Fünfschritt methodischen Handelns (Analyse der Rahmenbedingungen; Situations- bzw. Problemanalyse; Zielentwicklung; Planung, Operationalisierung; Evaluation). Die Aufgabe für Lehrende ist es nun, eine Verbindung zwischen all diesen verschiedenen Methoden-Begrifflichkeiten darzustellen. Demnach wäre der Ausgangspunkt die „Soziale Arbeit als Disziplin“. Diese nutzt die Methoden der Gemeinwesenarbeit, Gruppenarbeit und Einzelfallarbeit (nach Mennemann & Dummann, 2020). Das Vorgehen in diesen Methoden entspricht dem genannten Fünfschritt. Und zur Umsetzung der einzelnen Schritte werden wiederum Methoden der empirischen Sozialforschung benötigt, sowie Methoden der Gesprächsführung, Kenntnis um die Bedeutung von Lernbiografien, Theorien abweichenden Verhaltens, etc. Diesen Zusammenhang muss ich als Lehrende meinen Studierenden kommunizieren; ich kann nicht davon ausgehen, dass sie ihn von allein erkennen.

Dieses beinhaltet einen entsprechenden Perspektivenwechsel auf meine eigene Fachdisziplin Psychologie: Wenn Soziale Arbeit eine eigenständige Disziplin mit eigenen Methoden ist, dann sind alle anderen Disziplinen (unter anderem die Psychologie) ihre Hilfswissenschaften. Und das bedeutet zum Beispiel, dass ich nicht mehr die „Grundlagen der Psychologie als Wissenschaft“ lehre, sondern die „Grundlagen der Psychologie als Hilfswissenschaft für die Soziale Arbeit“. Ich lehre dann auch nicht mehr „Methoden und Techniken wissenschaftlichen Arbeitens“, sondern ich lehre „Methoden und Techniken wissenschaftlichen Arbeitens in der Sozialen Arbeit“. In der Quintessenz heißt dies nichts Anderes, als dass nicht mehr *die* Methoden wissenschaftlichen Arbeitens (einer bestimmten Fachdisziplin) im Fokus stehen, sondern eine Vermittlung von **Wissenschaftlichkeit als allgemeiner Haltung und Einstellung**.

⁶ David Pace (u.a. 2017) hat hier mit dem Modell des „Decoding the Disciplines“ einen praktikablen Ansatz für eine entsprechende Didaktik erarbeitet.

Fazit

Wenn wir als Lehrende, gleich welcher Wissenschaft angehörig, als Fachdisziplinen-übergreifenden, gemeinsamem Konsens also eine entsprechende *Grundhaltung* der *Wissenschaftlichkeit* festhalten können, bedeutet dies für uns als „Hilfswissenschaftler*innen“ in der Sozialen Arbeit eine Anerkennung der Sozialen Arbeit als eigenständige wissenschaftliche Disziplin. Für die Soziale Arbeit als Disziplin und als Wissenschaft bedeutet dies im Umkehrschluss, dass sie uns Vertreter*innen ihrer Hilfswissenschaften ihr Vorgehen und ihre Methoden erläutert und mitteilt, in welchen Bereichen sie unsere Hilfe und Unterstützung anfragt. In aller Kürze: Wir müssen alle mehr miteinander kommunizieren.

Prof. Dr. Alexandra Lehmann ist Diplom-Psychologin und hat seit 2011 eine Professur für Psychologie an der Ev. Hochschule Rheinland-Westfalen-Lippe im Fachbereich I (Soziale Arbeit, Bildung, Diakonie).

Literatur

- Bastian, J. & Groß-Mlynek, L. (2018). *Lerntechniken und Wissensmanagement* (2. Auflage). Wien: Huter und Roth.
- Bohl, T. (2008). *Wissenschaftliches Arbeiten im Studium der Pädagogik* (3. Auflage). Weinheim: Beltz.
- Brauner, D. J. & Vollmer, H.-U. (2008). *Erfolgreiches wissenschaftliches Arbeiten* (3. Auflage). Sternenfels: Verlag Wissenschaft & Praxis.
- Centeno García, A. & Kenneweg, C. (2019). Kritisch. Denken. Lernen. Einleitung zum Themenheft. *die hochschullehre. Interdisziplinäre Zeitschrift für Studium und Lehre*, S. 891-904. [Download unter: <http://www.hochschullehre.org/?p=1588>; 15.02.2021].
- Lambers, H. (2020). *Theorien der Sozialen Arbeit. Ein Kompendium und Vergleich* (5. Auflage). Opladen: Barbara Budrich.
- Mennemann, H. & Dummann, J. (2020). *Einführung in die Soziale Arbeit* (3. Auflage). Baden-Baden: Nomos.
- Müller, B. (2017). *Sozialpädagogisches Können. Ein Lehrbuch zur multiperspektivischen Fallarbeit* (8. Auflage). Freiburg i. Br.: Lambertus.
- Pace, D. (2017). *The Decoding the Disciplines Paradigm. Seven Steps top Increased Student Learning*. Bloomington: Indiana University Press.
- Schilling, J. (2020). *Didaktik/ Methodik Sozialer Arbeit* (8. Auflage). München: Ernst Reinhardt.
- Schrader, J. (2008). *Lerntypen bei Erwachsenen. Empirische Analysen zum Lernen und Lehren in der beruflichen Weiterbildung* (2. Auflage). Bad Heilbrunn: Julius Klinkhardt.
- Wahl, D. (2020). *Wirkungsvoll unterrichten in Schule, Hochschule und Erwachsenenbildung*. Bad Heilbrunn: Julius Klinkhardt.

Joachim König, Sebastian Ottmann & Alexandra Richter

Wirkungen im Blick, aus verschiedenen Perspektiven – ein kurzes Fazit

Mit einem zusammenfassenden Blick auf den anregenden und konstruktiven Dialog auf Augenhöhe zwischen Vertreter*innen der Praxis, der Verbände und der Politik einerseits sowie Akteur*innen aus Forschung, Lehre und Sozialwissenschaften im Rahmen der insgesamt elf Beiträge andererseits, wurde die große Bandbreite der inzwischen differenzierten empirisch-methodischen, praktisch-konzeptionellen sowie der strategischen und sozialpolitischen Debatten zu diesem Themenkreis deutlich.

Die hier formulierten zentralen Erkenntnisse versuchen diese Debatte im Rahmen der Konferenz zu bündeln und zusammenzufassen:

Die Erstellung von Wirkungsindikatoren-Datenbanken erscheint für die Felder der Sozialen Arbeit als ein grundsätzlich lohnenswertes Unterfangen, das allerdings mit einem enormen Arbeitsaufwand verbunden ist und auch eine enge und intensive Zusammenarbeit vieler Akteur*innen aus ganz unterschiedlichen Bereichen voraussetzt. Der große Benefit einer solchen – frei zugänglichen – Datenbank bestünde nicht nur darin, dass Wirkungsanalysen in den verschiedenen Feldern vor dem Hintergrund bereits formulierter ‚Indikatoren-Sets‘ erstellt werden könnten und so mit einem wesentlich geringeren methodischen Vorbereitungs- und empirischen Planungsaufwand verbunden wären. Sondern auch die für die Praxisfelder wichtigen inhaltlichen und konzeptionellen Debatten, um Effekte und Wirksamkeit von Maßnahmen und auch um sinnvolle Konsequenzen daraus könnten, unterfüttert mit einer wesentlich breiteren inhaltlichen Substanz, geführt werden.

Die evidenzbasierte, also Kennzahlen voraussetzende, Steuerung einer Praxis bietet große Vorteile im Bereich der Begründung und Legitimierung von Maßnahmen und bezogen auf die Anhaltspunkte für die ständige Verbesserung ihrer Qualität. Jedoch gibt es gerade hier legitime kritische Anfragen an eine deutlich zunehmende Ökonomisierung der Sozialen Arbeit insgesamt. Vor allem mit Blick auf die Finanzierung von Leistungen und dem zwangsläufig damit verbundenen Prinzip des ‚pay by results‘ wird die Gefahr deutlich, die Praxis der Sozialen Arbeit auf ihre Rolle als Investitionsobjekt in einem ‚Sozialinvestitionsstaat‘ zu reduzieren und die Bedeutung des Sozialstaates als ‚Gewähr-Leister‘ der Daseinsvorsorge damit immer weiter in den Hintergrund zu drängen.

Zwei Herausforderungen haben sich für einen künftigen, verantwortungsvollen Umgang mit diesem Dilemma – auch vor dem Hintergrund der Frage, wie von den Verantwortlichen vermutete oder gefundene Wirkungseffekte priorisiert werden sollen und wer die dazu notwendige Definitionsmacht besitzt – als wichtig erwiesen:

- Die Klärung der Werthintergründe und normativen Entscheidungen bei der Erstellung von Wirkmodellen und eine Offenlegung möglicherweise unterschiedlicher Sozialstaats-

Verständnisse unter den Beteiligten im Rahmen von Wirkungsanalysen und im Austausch mit denjenigen, die im Anschluss an den Entscheidungen um die Verwertung der Befunde beteiligt sind.

- Die Beteiligung der Fachkräfte und – soweit möglich – auch der Klient*innen im empirischen Prozess einer Wirkungsanalyse, nicht nur, um dadurch die Validität der Befunde zu optimieren, sondern auch als Ausdruck einer grundsätzlich an dem Grundsatz der Partizipation orientierten und interessierten Vorgehensweise.

Aus der Erfahrung von durchgeführten Projekten zeigte sich, dass es auch alternative Verfahren gibt, um Effekte und Wirkungen aufzuzeigen. Hier wurde der Ansatz der theoriebasierten Evaluation anhand eines Projektes vorgestellt. Mit diesem Ansatz ist es möglich eine Einschätzung zu treffen, welchen Anteil das untersuchte Projekt an Effekten hat. In einem weiteren Projekt wurde auf die Thematik Wirkmodelle eingegangen und dargestellt, wie diese in der Bewährungshilfe in der Schweiz entwickelt wurden. Hier zeigte sich, dass Wirkmodelle eine wichtige Grundlage für darauf aufbauende Wirkungsanalyse sind und dadurch dann neben der konkreten Wirkung auch Wirkfaktoren identifiziert werden können. Es zeigte sich auch, dass die Evaluationsaktivitäten in den Organisationen der Sozialen Arbeit sehr unterschiedlich ausgeprägt sind. Hier wurden Ergebnisse aus der Schweiz vorgestellt. In der Zukunft erscheint es sinnvoll, wenn sich Einrichtungen und Träger dem Thema Wirkungsorientierung und Wirkungsanalyse widmen möchten, dass diese auch ihre Ressourcen und Kapazitäten im Bereich der Evaluation ausbauen.

An der Schnittstelle von Wissenschaft und Praxis stoßen zuweilen unterschiedliche Denklogiken aufeinander. Dies soll anhand zweier Beispiele verdeutlicht werden:

- Wissenschaftliche Begleitforschung läuft immer zwischen zwei möglichen Gegensätzen. So soll die Forschung einerseits möglichst nach den Gütekriterien Objektivität, Reliabilität und Validität erfolgen. Aber andererseits muss sie auch in den Alltag der Praxis passen. Diesem Dilemma kann unter Umständen nur mit „methodischen Kompromissen“ entgegnet werden – und natürlich unter Offenlegung der Abweichungen und unter Berücksichtigung der Regeln der guten wissenschaftlichen Praxis.
- Die Quantifizierung von in der Praxis relevanten Ereignissen bedarf häufig einer qualitativ ergänzenden Forschungsmethodik. Gerade in der wissenschaftlichen Begleitforschung des Praxisalltags der Sozialen Arbeit wird deutlich, dass das Messbarmachen von sozialem Handeln nicht immer ausreichend Erklärungen über das „Warum“ der Handlungen liefert. Ein methodisch begründeter Method-Mix ermöglicht hingegen die Validierung quantifizierter Ergebnisse durch eine ergänzende qualitative Auswertung und ermöglicht so die Berücksichtigung „weicher“ Faktoren.

Das Wissen um diese unterschiedlichen Handlungs- oder Denklogiken ermöglicht dennoch eine zielführende, wissenschaftlich fundierte Begleitforschung der Praxis in der Sozialen Arbeit. Essentiell ist es, sich dieses Wissen bewusst zu machen und es auch zum Thema machen. So kann die Herausforderung der Wirkungsmessung unter Berücksichtigung der vielen Besonderheiten der

sozialen Praxis dennoch Rechnung getragen werden. Schlussendlich kann konstatiert werden, dass es zuweilen eines pragmatischen Vorgehens bedarf und dieses durch einen rekursiven Prozess des Aushandelns zwischen der reinen Lehre der Forschung einerseits und dem Anerkennen der sozialen Praxis andererseits gekennzeichnet ist.