

4 Wirkungsmessung in der Sozialen Arbeit

Nicht mehr die Anzahl der Betreuten, sondern ihre gesteigerte Lebensqualität steht im Mittelpunkt.

Dieser Satz macht die wirkungsorientierte Denkweise deutlich. Wie in Kapitel 3.2.1, Wirkungsdebatte, bereits beschrieben, ist es aktuell schwierig, sich in der Sozialen Arbeit nicht mit Wirkungsmessung auseinander zu setzen. Es stellt sich dabei die Frage, wie diese Auseinandersetzung sinnvollerweise geschehen kann. Somit ist es notwendig, eine Strategie zu entwickeln, um die Wirksamkeit für Stakeholder*innen und Leistungsträger*innen darzustellen und gleichzeitig die Grundwerte der Sozialen Arbeit, wie z. B. Partizipation der Klient*innen, Ergebnisoffenheit und Professionalität der Fachkräfte nicht einzuschränken. Parallel ist im Sinne der realistischen Evaluation zu beachten, dass der Kontext mit einbezogen wird.

Die Ansätze der Wirkungsmessung können grob in verschiedene Gruppen aufgeteilt werden:

- Einfache Wirkungsanalysen: erläutern selten die Wirkmechanismen und zeigen nur einzelne Wirkungen auf
- Umfassende Wirkungsanalysen: beinhalten mehrere Wirkungen und Ebenen und u. U. Quantifizierungen (z. B.: SROI)
- „Instrumente zur Berichtslegung (z. B.: Social Reporting Standard)
- Instrumente zum Rating (z. B.: IRIS, GRI)
- Strategieinstrumente (z. B.: Wirkungsmodell, Theory of Change)
- Steuerungsinstrumente (z. B.: wirkungsorientierte Steuerungsbox; Logical Framework“ (Grünhaus und Rauscher 2022, S. 514–515)

In diesem Kapitel werden die Ergebnisse der systemischen Literaturanalyse dargestellt. Es wurden dabei Inhalte zu den Bereichen umfassende Wirkungsanalyse, Strategieinstrument und Steuerungsinstrumente berücksichtigt. Auf einfache Wirkungsanalysen, Instrumente zur Berichtslegung und zum Rating wurde nicht eingegangen, da sie für die Beantwortung der Forschungsfrage als nicht relevant bewertet wurden.

Zunächst wird der Zweck von Wirkungsmessung beschrieben, nachfolgend wird auf die verschiedenen Methoden eingegangen, welche im Rahmen der realistischen Evaluation in die Praxis übertragen werden können. Dieses Kapitel teilt sich in die Beschreibung des Zweckes und der Grenzen sowie der Methoden, Kategorisierung von Wirkungsmessung, Wirkmodelle, Wirkungsanalysen, Stakeholder*innenperspektiven und Indikatorenbildung auf. Anschließend werden wirkungsorientierte Steuerungsmodelle vorgestellt, bevor auf die Implementierung von Wirkungsmessung und die vermeintlichen Folgen eingegangen wird.

4.1 Zweck und Grenzen der Wirkungsmessung

Es gibt unterschiedliche Zwecke, für die Wirkungsmessung genutzt werden kann. Im Folgenden erfolgt eine Aufzählung, die keinen Anspruch auf Vollständigkeit hat, die aber die in der Literatur häufig genannten Gründe beinhaltet. Dabei wird vorausgesetzt, dass die Messung an die jeweiligen Gegebenheiten und Ressourcen der NPO angepasst ist.

Ein wichtiger Zweck, der durch Wirkungsmessung erfüllt werden kann, ist die Missionserfüllung und die damit verbundene Steigerung des gesellschaftlichen Mehrwertes. Damit einher gehen eine hohe Transparenz und die Darstellung der Organisationsziele sowie des Leistungsportfolios. NPOs können dadurch eine eigene „Marke“ entwickeln und so z. B. Alleinstellungsmerkmale hervorheben. Durch die Auseinandersetzung mit der eigenen Mission und den Zielen erfolgt eine Reflexion der eigenen Wirkweise, welche z. B. in Kooperationen mit und in Abgrenzung von anderen Partner*innen hilfreich sein kann. Wird die Wirkungsmessung in Form von Wirkungsorientierung mit dem strategischen Management verknüpft, dient sie zur innovativen Steuerung der Organisation als Alternative zum

Ansatz der Input-/Outputsteuerung. Gut umgesetzt unterstützt sie eine effizientere und effektivere Mittelverwendung und führt dazu, dass die positiven Wirkungen bewusst gemacht werden. Häufig wird Wirkungsmessung als Legitimation für Finanzgeber*innen verwendet und zur Außendarstellung des Erfolges bzw. um der Nachweispflicht nachzukommen. Ist sie valide und evidenzbasiert, kann sie als Datengrundlage für politische Entscheidungen dienen (s. Umsetzung in England) oder zur fachwissenschaftlichen Debatte beitragen. (Grünhaus und Rauscher 2022, 521ff; Kehl et al. 2018, S. 277; Pree 2019, S. 48) Sie kann „zivilgesellschaftliche Infrastruktur und eine bürgerschaftliche Kultur im lokalen Nahraum mitentwickeln – und damit, ökonomisch gesprochen, an der Produktion von Dienstleistungen und (Kollektiv-) Gütern mitwirken, von denen niemand ausgeschlossen werden kann“ (Kehl et al. 2018, 288f).

Die Grenzen der Wirkungsmessung sind dagegen u. a. Folgende:

Wirkungen von Sozialer Arbeit sind i. d. R. zeitlich versetzt, während parallel das uno-actu-Prinzip gilt. Beide Gegebenheiten und die Kombination der beiden erschweren die Messung von Wirkungen. Z. B. ist eine Fachkraft in der Suchtberatungsstelle auf die Mitarbeit der*des Adressaten*in angewiesen. Ihr „Produkt“, die Beratung, wird bei der Herstellung „konsumiert“. Ob die Beratung dann dazu geführt hat, dass die*der Adressat*in später den Drogenkonsum beendet hat, ist zum Ende der Beratung häufig (noch) nicht messbar. Darüber hinaus ist eine Objektivierbarkeit von Wirkungen schwer umsetzbar. In Kapitel 4.2.5 wird bei der Indikatorenbildung auf den Punkt im Detail eingegangen. Die begrenzten Ressourcen sowie der Aufwand für die Umsetzung von Wirkungsmessung sind ebenfalls als Grenzen zu benennen. (Pree 2019, S. 48)

Wirkungsmessung sollte nicht als Selbstzweck dienen, sondern für ein bestimmtes Ziel genutzt werden. Dieses Ziel lässt sich in vier Funktionen einordnen. Diese sind der nachfolgenden Grafik zu entnehmen. Es handelt sich dabei um die Erkenntnis-, Kontroll-, Entwicklungs- und Legitimationsfunktion. (Baumgartner und Haunberger 2023, S. 30) Diese Funktionen entsprechen auch Leitungsaufgaben, die von Führungskräften umgesetzt werden müssen. Im Rahmen der wirkungsorientierten Steuerung kann der Nutzen der Wirkungsmessung mit den Managementaufgaben kombiniert werden, s. Kapitel 4.3.

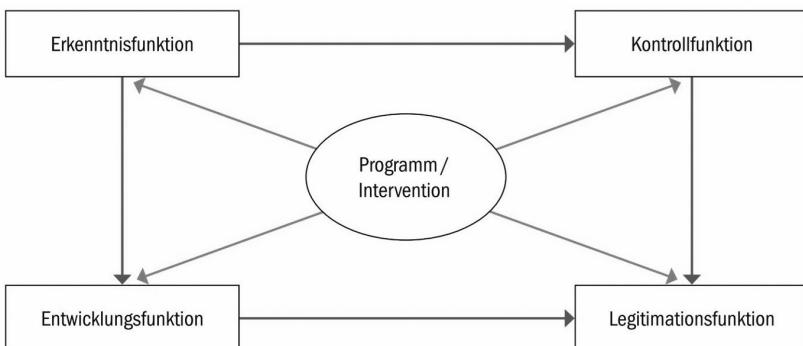


Abbildung 5: Zielfunktionen von Wirkungsevaluation (in Anlehnung an Stockmann 2004) (ebd.)

Je nach Zweck-Einordnung können die Ansätze und die Umsetzung der Analyse differieren. Darauf wird in Kapitel 4.2.3 eingegangen.

4.2 Methoden der Wirkungsmessung

In der Literatur werden zahlreiche Methoden für die Wirkungsmessung dargestellt. Es ist dabei zwischen verschiedenen Arten der Instrumente zu unterscheiden. Einige, z. B. der Social Return on Investment (SROI) oder die Social Impact Bonds (Grünhaus und Rauscher 2021, S. 3), legen den Fokus auf den monetären Wert der Sozialen Arbeit. Es wird dabei der finanzielle Mehrwert sozialer Dienstleistung errechnet. Darüber hinaus gibt es Instrumente wie die Balanced Scorecard (BSC) von Kaplan und Norton oder die Wirkungstabelle der International Group of Controlling (IGoC), welche von Instrumenten aus der Wirtschaft abgeleitet sind und auf die Gegebenheiten der Sozialen Arbeit angepasst wurden. Sie beziehen verschiedene Stakeholder*innenperspektiven mit ein und legen den Schwerpunkt auf inhaltliche Wirkungen. Darunter gibt es Ansätze, die Aspekte der realistischen Evaluation inkludieren und Kontextfaktoren mit einbinden, wie z. B. die Wirkungsanalyse mit dem Wirkungsradar nach Ottmann. (2023, 50f)

Im CSI-Gutachten werden die Ansätze verglichen und auf die Anwendbarkeit in Wohlfahrtsverbänden überprüft. Als Grundlage dieser Arbeit

wurde die Tabelle der CSI analysiert. Rein-monetäre Ansätze wurden nicht berücksichtigt, da mit Blick auf die professionstheoretische Diskussion (s. Kapitel 3.2.3) der Fokus auf die qualitativen Wirkungen der Sozialen Arbeit und nicht auf die finanziellen Aspekte gerichtet werden soll. Darüber hinaus wurden alle Ansätze ignoriert, die nicht als Wirkungsmessung beurteilt wurden.

Das Transparenz-Gutachten hat folgende Ergebnisse gezeigt:

- 1) je nach Wirkungsdimension, Interventionsfeld und -form sind unterschiedliche Modelle und Methoden anzuwenden
- 2) Wirkmodelle müssen zum Einsatz kommen und entsprechend der Zielsetzung der Wirkungsmessung ausgewählt werden
- 3) es gibt bisher keine Verfahren, die „die vieldimensionalen Wirkungen sozialer Arbeit in Zahlen, also quantitativ [...] bestimmen, schon gar nicht in Euro und Cent.“ (Kehl et al. 2016, S. 5)

Sowohl im Gutachten als auch in der weiterführenden Literatur wird deutlich, dass aktuell kein Ansatz vollumfängliche Ergebnisse liefert (s. Kapitel 3.2.2). Daher ist eine Kombination verschiedener Methoden und Ansätze notwendig. (ebd., 54f)

Wie in Kapitel 2.3 bereits beschrieben, werden als Grundlage für die verschiedenen Instrumente von wirkungsorientierter Steuerung Wirkungsmodelle genutzt. Sie dienen der Komplexitätsreduzierung und helfen, eine Übersicht über die Maßnahmen zu verschaffen. (Kränzl-Nagl et al. 2019, S. 34) Wirkmodelle haben in mehrfacher Hinsicht einen Nutzen. Sie dienen der transparenten Darstellung von Wirkmaßnahmen z. B. für Leistungsträger*innen und zur Legitimation der Arbeit. Während der Erstellung findet eine Reflexion der konzeptionellen und fachlichen Arbeit statt, welches u. a. zu einer Weiterentwicklung der Maßnahme führt. Zusätzlich sind sie eine Grundlage für die Darstellung der Sozialen Arbeit in der Gesellschaft und fördern die berufliche und professionelle Identität der Fachkräfte. Darüber hinaus dienen sie als Basis für Wirkungsanalysen. (Ottmann und König 2023, 47ff, 70) Außerdem unterstützen sie bei der in- und externen Kommunikation von Maßnahmen und können den Planungs-

prozess strukturieren. Wirkungsmodelle erleichtern die Fokussierung auf die wichtigen Elemente der Maßnahme und die Berichterstattung darüber. Bei der Einarbeitung von neuen Mitarbeitenden können sie genutzt werden und sie dienen der Steuerung von Bewertungs- und Evaluationsprozessen. (Baumgartner und Haunberger 2023, 117f) Im Wesentlichen unterscheiden sich die Modelle in der Darstellung und dem Umfang. In Kapitel 4.2.2 werden verschiedene Modelle erläutert. Die Grundelemente eines Wirkmodells werden mit der IOOI-Logik (s. Abb. 6) beschrieben. IOOI bedeutet: Input (Ressourcen), Output (Maßnahmen), Outcome (Wirkungen bei der Zielgruppe) und Impact (gesellschaftliche Wirkung). Ab der Stufe des Outcomes wird von Wirkungen gesprochen.

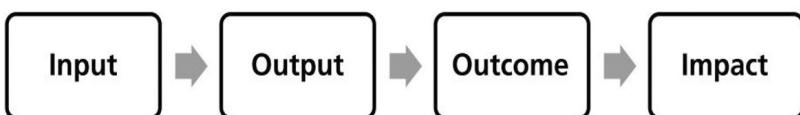


Abbildung 6: IOOI Logik (Ottmann und König 2023, S. 21)

Die im Bild dargestellte Wirkungskette beschreibt einen groben Ablauf. Sie existiert in dieser Form in der Praxis nicht, da, wie schon erwähnt, eine lückenlose Kausalität zwischen Leistung und erzielten Wirkungen aufgrund von externen Faktoren nicht nachgewiesen werden kann. Diese Lücken lassen sich durch aufwendige empirische Studien, Handlungsmodelle oder Wirkungsplausibilisierung auffangen. (Burmester 2020, S. 43)

Bei einem systemischen Blick auf die Kausalität von Ursachen und Wirkungen sind folgende Punkte zu berücksichtigen: verschiedene sich untereinander bedingende Faktoren, welche sowohl systemintern als auch -extern sind, beeinflussen die Wirkungen. Sie sind daher nicht durch einzelne Faktoren zu beschreiben, sondern nur im Rahmen des Kontextes, in dem sie sich entfalten. „Ursachen und Wirkungen lassen sich in komplexen Interventionen nicht eindeutig identifizieren und voneinander trennen“ (Baumgartner und Haunberger 2023, S. 41). Durch Rückkopplungseffekte können Wirkungen zu Ursachen werden und umgekehrt. „Kausalität ist kontext- und zeitgebunden, daher kann dasselbe Ereignis unter anderen Bedingungen oder zu einem anderen Zeitpunkt unterschiedliche Folgen haben“ (ebd.).

Aus diesem Grund wird das Modell der Wirkungskette bei Pawson und Tilley um den Kontext ergänzt (s. Abb. 4, Seite 50). Diese Grundlagen sollten durch einen zeitlichen Faktor, welcher beschreibt, wann die Wirkungen erwartet werden und die Persistenz (Nachhaltigkeit) sowie den Transfer auf andere Kontexte ergänzt werden. Die Nachhaltigkeit kann in drei Kategorien eingeteilt werden: stabile Wirksamkeit, Wirksamkeitsgewinn, Wirksamkeitsverlust. (Ottmann und König 2023, 42ff)

Werden die äußeren Einflüsse bereits im Wirkmodell detailliert erfasst, können sie später in der Analyse ausgewertet werden. Beim Transfer der Maßnahme kann zusätzlich frühzeitig eingeschätzt werden, ob die gleichen Wirkungen auch in einem anderen Kontext zu erwarten sind. Es wird zwischen vier verschiedenen Arten von Kontextfaktoren unterschieden: Strukturen (der Einrichtung/Träger), Inputs (Mittel des Leistungsträgers, die einfließen, z. B. VK-Anteile), Incomes (Ressourcen der Klient*innen und des Umfeldes), Rahmenbedingungen (z. B. gesellschaftliche, rechtliche oder sozialpolitische Kontexte) (Beywl und Niestroj 2009, S. 139).

Für die Beschreibung der Außeneinflüsse bietet sich eine Matrix (s. Abb. 7) an. Mithilfe dieser Matrix können die einzelnen Einflüsse aufgelistet sowie bzgl. der Größe des zu erwartenden Einflusses und ob es sich um einen positiven oder negativen Einfluss handelt, bewertet werden.

Außen- einfluss	Beschrei- bung	Einfluss pos./neg.	Einfluss auf Klient*in			Einfluss auf Maßnahme		
			gering	stark	gering	stark		
			1	2	3	4	5	6

Abbildung 7: „Matrix zur Identifikation und Beschreibung von Außeneinflüssen“ (Ottmann und König 2023, S. 52)

Im Rahmen der Erstellung eines Wirkmodells werden die Wirkungen in Haupt- und Teilwirkungen aufgeteilt. Durch diese Aufteilung lassen sich Wirkmechanismen erkennen, wodurch einzelne Methoden besser

abgestimmt werden können und die Wahrscheinlichkeit, dass das Angebot die gewünschte Wirkung erzielt, erhöht wird. (ebd., S. 60) Die Wirkmechanismen werden wiederum in zwei Arten unterschieden: dem sozialen Mechanismus, welcher die Ebene der Klient*innen (z. B. Verhaltensänderungen) beschreibt und dem Prozessmechanismus, der sich auf den gesamten Ablauf des Angebotes (Metaebene) bezieht. (ebd., S. 61)

Zur Nutzung von Synergieeffekten ist es hilfreich, aus einzelnen Wirkmodellen Ankerwirkmodelle zu erstellen. Diese haben eine gewisse Allgemeingültigkeit für gleiche bzw. ähnliche Arbeitsbereiche (z. B. Erziehungsberatungsstelle der Diakonie und der Caritas oder bei gleichen Aufgabengebieten an verschiedenen Standorten eines Trägers). (ebd., S. 72)

Wie bereits erwähnt, ist das Wirkmodell ein Hauptbestandteil der Wirkungsanalysen. Als Wirkungsanalyse⁹ wird die Messung von intendierten, konzeptionell geplanten und nicht-intendierten Wirkungen, welche während der Durchführung entstehen, beschrieben. Je nach Quelle unterscheiden sich der Abläufe der Wirkungsanalysen. Sie bestehen i. d. R. aus einer Planungs-, einer Umsetzungs- und einer Evaluationsphase. Je nach Autor*in erfolgen diese Phasen in unterschiedlich vielen Schritten. Eine detaillierte Beschreibung der verschiedenen Ansätze erfolgt in Kapitel 4.2.3.

Im Rahmen einer Wirkungsanalyse erfolgt neben der Stakeholder*innenanalyse die Indikatorenbildung. Dabei gilt zu beachten, dass Indikatoren immer normativ sind und daher einer nachvollziehbaren und angemessenen Beschreibung bedürfen, damit sie verlässlich und zielführend eingesetzt und gegenüber Kritik argumentiert werden können. (Ziegler 2017, 37f)

Wirkungsevaluation wird durchgeführt, um Wirkmechanismen zu erklären. Ziel ist es, unter Berücksichtigung der Kontextbedingungen Ursachen für Wirkung zu hinterfragen. Sie beantwortet damit die Frage: Was führt dazu, dass es funktioniert? Basis der Wirkungsanalyse ist die Auswertung der Interaktion von Ressourcen, professionellen Entscheidungen und Kontextbedingungen. Um tragfähige Ergebnisse zu ermit-

9 Das Wort Wirkungsanalyse wird in dieser Arbeit synonym mit dem Begriff Wirkungsevaluation genutzt.

teln, bedarf es dabei Wiederholungen über einen längeren Zeitraum und Verifikation durch Triangulation. (Ziegler 2017, 43f)

Wirkungen werden auf verschiedenen Ebenen gemessen. Je nach Ausmaß und Ebene der Auswirkungen werden sie unterschiedlich eingestuft. In der nachfolgenden Abb. 8 sind die verschiedenen Wirkungs-Ebenen zu erkennen. Die unteren drei Stufen gehören der Kategorie Output an. Damit sind Maßnahmen und Angebote bzw. Aktivitäten gemeint. Ab der vierten Stufe wird von Outcome gesprochen, diese beziehen sich auf die Wirkungen bei der Zielgruppe. Ab Stufe sieben wird der Impact beschrieben. Dieser bezieht sich auf die Auswirkungen für die Gesellschaft.

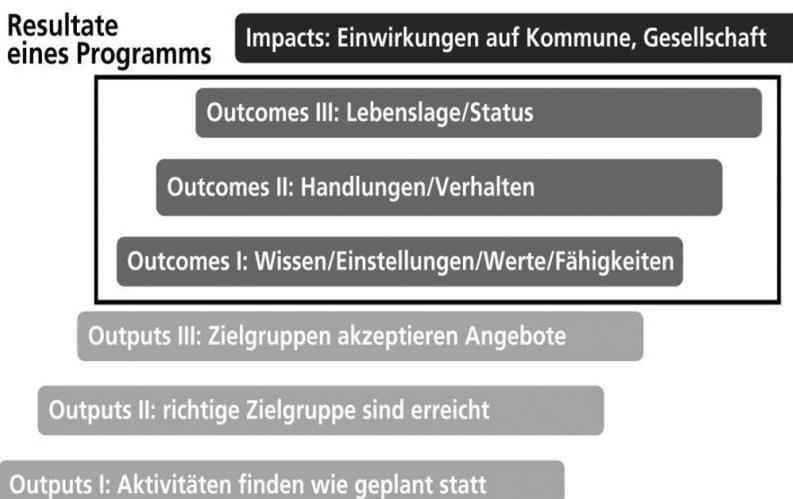


Abbildung 8: Resultate eines Programmes (Ottmann und König 2023, S. 59)

In den folgenden Kapiteln wird zunächst eine Kategorisierung der Ansätze von Wirkungsmessung erläutert, anschließend werden verschiedene Wirkmodelle vorgestellt. In Kapitel 4.2.3 werden Wirkungsanalysen beschrieben, bevor auf die Indikatorenbildung und Stakeholder*innenperspektiven eingegangen wird.

4.2.1 Kategorisierung der Ansätze

Unterschiedliche Ansätze der Wirkungsmessung wurden in den vorherigen Kapiteln bereits genannt. Diese Ansätze lassen sich in verschiedene Kategorien einteilen.

Wirkungsmessung kann im Rahmen der Evaluationsforschung zum Einsatz kommen. In dem Zusammenhang werden objektivierte Verfahren genutzt, um Sachverhalte zu bewerten. Es wird dabei auch analysiert, wie Programme/Projekte funktionieren.

Eine weitere Kategorie ist die der sozialen Rechenschaftslegung (oben auch Legitimation genannt). In dem Kontext wird Wirkungsmessung genutzt, um soziale Wirkungen finanziell messbar zu machen, z. B. beim SROI, Cost-Utility-Analysen, Cost-Benefit-Analysen. Dabei geht es nicht um eine sozialwissenschaftliche Erläuterung der Prozesse, sondern um die Darstellung der Effekte als soziale Investition. (Schober und Rauscher 2014, 265ff) In den anglo-amerikanischen Ländern lag der Fokus zu Beginn der Wirkungsdiskussion auf der Darstellung der monetären Perspektive, um den Mitteleinsatz zu legitimieren. In den letzten Jahren wurde von der Vorgehensweise zunehmend abgewichen, da eine verlässliche Monetarisierung der Wirkungen nicht umsetzbar ist. Der Fokus liegt daher aktuell auf der Darstellung von einer Bandbreite an quantitativen und qualitativen Ergebnissen. (Kehl et al. 2018, 290f)

Der dritte Bereich, in dem Wirkungsmessungen zum Einsatz kommen, ist der des Social Impact Assessment. Ursprünglich kommt diese Kategorie aus den USA als Umweltverträglichkeitsprüfung. In der Folge der Übertragung in die EU wurde der soziale Impact ergänzt. In diesem Rahmen werden die Auswirkungen für die natürliche und soziale Umwelt analysiert. In dem Zusammenhang werden der quality adjusted life years-Ansatz (QALY) sowie der Happy Planet Index genannt, wobei sie nicht als klassische Wirkungsanalysen angesehen werden. (Schober und Rauscher 2014, 265ff)

Im folgenden Kapitel werden die Wirkmodelle näher erläutert.

4.2.2 Wirkmodelle

Wirkmodelle dienen zur „Formulierung von Annahmen kausaler Wirkungszusammenhänge“ (Kehl et al. 2018, S. 280). Ein umfassendes Wirkmodell beinhaltet verschiedene Ebenen, Dimensionen und Stakeholder*innenperspektiven sowie die verschiedenen Maßnahmen. In den meisten Wirkmodellen wird mit Wirkungsketten gearbeitet und die Wirkungen werden auf verschiedenen Stufen dargestellt. „Das Wirkungsmodell ist das konzeptionelle Zentrum jeder Wirkungsanalyse und gibt gleichzeitig die grundlegende Wirkungslogik der betrachteten NPO wieder.“ (Grünhaus und Rauscher 2022, S. 515–517)

Es gibt zahlreiche Versionen von Wirkungsmodellen. In der nachfolgenden Grafik ist eine Übersicht zu sehen, welche eine grobe Zuordnung ermöglicht. Es wird dort zwischen verschiedenen Eingruppierungen wie bspw. der Aussagekraft, der Interventionsebene, des Grades des Theoriebezuges etc. unterschieden. Manche Modelle sind nicht klar zuzuordnen, da sie z. B. vor, während oder nach einer Maßnahme erstellt, werden können oder mehrere Ebenen (Makro-, Meso-, Mikroebene) abdecken, wie z. B. die Wirkungsbox von Grünhaus und Rauscher. Dennoch dient diese Tabelle zur Orientierung und bei der Unterstützung der Auswahl der Instrumente.

Wirkungsmessung in der Sozialen Arbeit

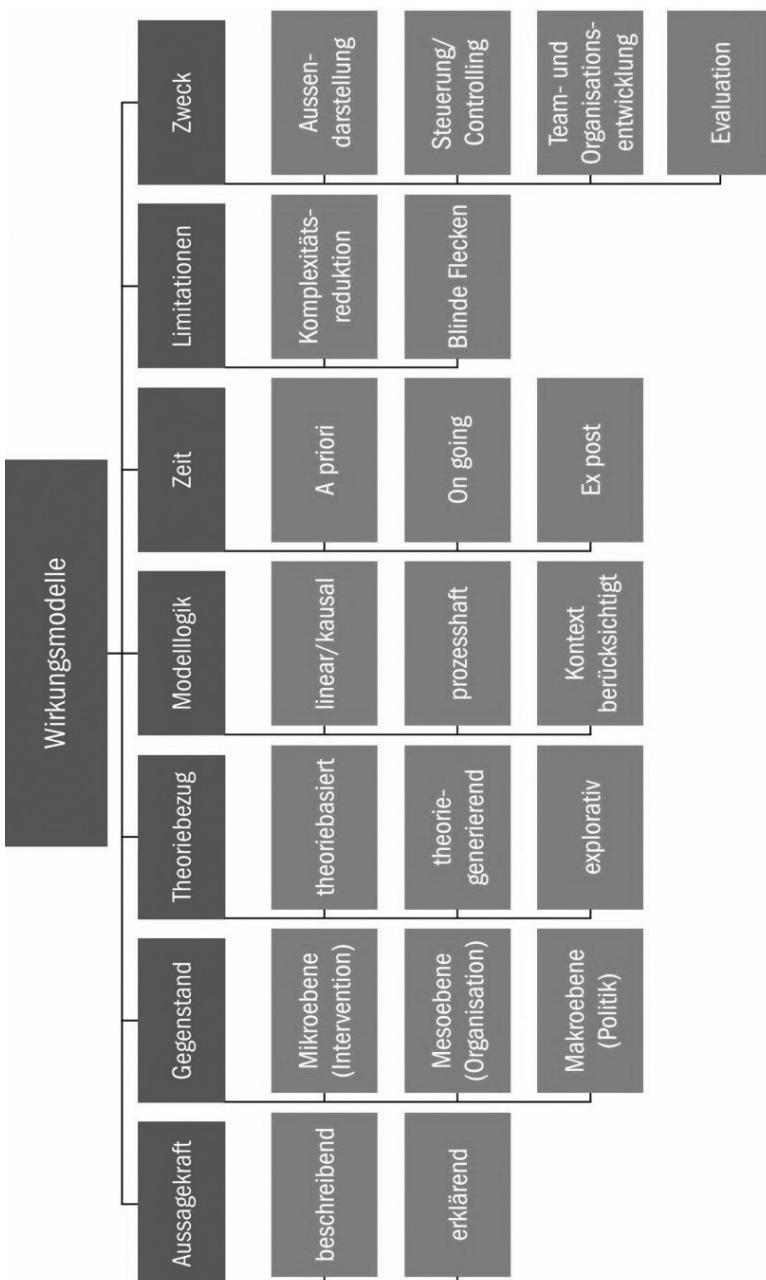


Abbildung 9: Auslegeordnung Wirkungsmodelle (Baumgartner und Haunberger 2023, S. 104)

In diesem Kapitel wird eine Auswahl der am häufigsten zitierten Wirkungsmodelle dargestellt.

Wirkungsmodell von Univation

Beywl und Niestroj prägen den Ansatz von Univation. Sie arbeiten mit einem Programmbaum und einer Resultattreppe. Der Programmbaum dient dazu, Maßnahmen zu evaluieren, in dem die Elemente der Wirklogik dargestellt und analysiert werden. Er hilft dabei, Programme besser zu verstehen und zukünftige Programme strukturiert planen zu können. Er basiert u. a. auf dem von Daniel Stufflebeam entwickelten CIPP-Modell. CIPP steht für die Begriffe Context, Input, Process, Product. Beywl und Niestroj differenzieren dabei das „C“, den Kontext, in die vier Bedingungen: Kontext, Income, Input und Struktur. (Beywl und Niestroj 2009, 140f) In Abb. 10 ist der Programmbaum zu sehen. Er ist in die vier Hauptbereiche Bedingungen, Plan, Umsetzung und Resultate unterteilt. Der Bereich der Resultate gliedert sich ebenfalls in verschiedene Kategorien auf. Auf diese wird bei der Erläuterung der Resultattreppe näher eingegangen. In der Grafik ist zu sehen, dass es sich beim Programmbaum um einen zirkulären Vorgang handelt. Dieses zeigt, dass die Ergebnisse der Evaluation Auswirkungen auf die Planung von neuen Programmen haben.

Wie oben erwähnt, nutzen Beywl und Niestroj, wie Ottmann und König, die Resultattreppe, welche bereits in Kapitel 4.2 auf Seite 53 dargestellt wurde, um die Wirkungen von Programmen detailliert zu beschreiben. Durch die Darstellung als Treppe wird deutlich, dass die verschiedenen Ergebnisse aufeinander aufbauen und die unteren Stufen abgeschlossen sein müssen, um die höheren zu erreichen. Die Einteilung der Stufen hilft zusätzlich dabei, dass die Programmverantwortlichen die Ziele ausführlich differenzieren müssen. Ab der Stufe vier sprechen Beywl und Niestroj von Wirkungen. (ebd., S. 144)

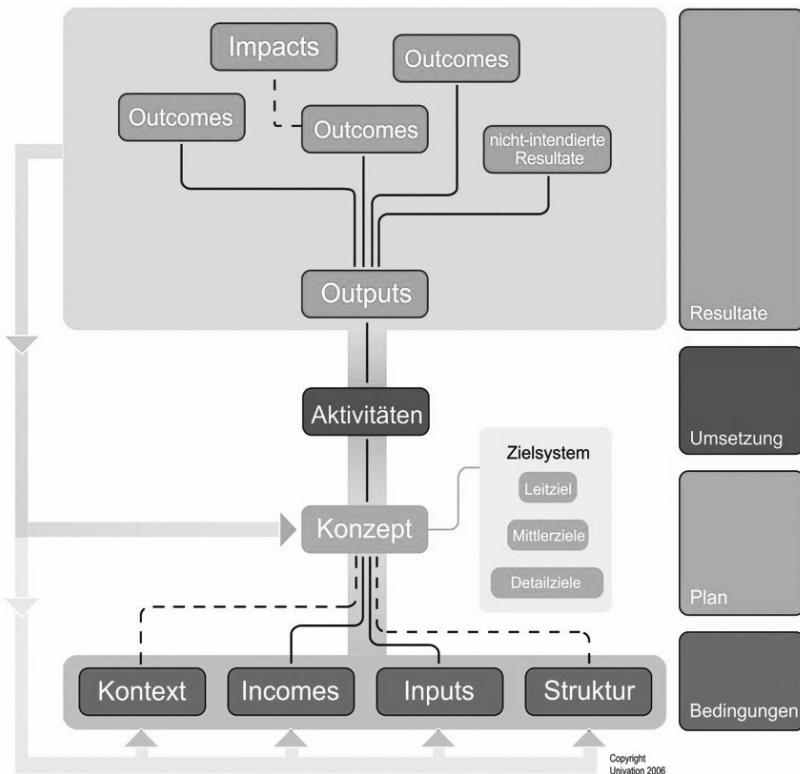


Abbildung 10: Programmbaum von Univation (Beywl und Niestroj 2009, S. 142)

Vorgehensmodell zur Wirkungsmessung der Stiftung ZEWO

Die Stiftung Zewo bedient sich zur Wirkungsmessung verschiedener Wirkungsmodelle, z. B. des Logical Framework Modells, des Logic Modells, des CIPP oder des Programmbaums. Sie hat eine Vorgehensweise zur Wirkungsmessung daraus entwickelt. Ziel ist die Steigerung von Effektivität und Effizienz sowie Lernen, Lenken und Legitimieren (s. Abb. 11).

Zewo orientiert sich für die Wirkungsmessung an den UN Millennium Development Goals (MDGs), den Grundsätzen der internationalen Entwicklungszusammenarbeit und den Menschenrechten. Schwerpunkt des Modells ist die Partizipation der Zielgruppe. Weitere Stakeholder*innen werden nicht explizit involviert.

Methoden der Wirkungsmessung

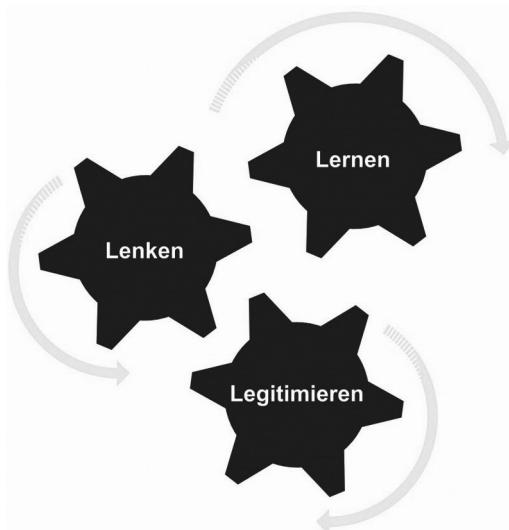


Abbildung 11: Gründe für Wirkungsmessung in NPO
(Stiftung Zewo 2013)

Wie auch bei anderen Wirkungsmodellen liegt die IOOI-Logik dem Modell zugrunde. In der nachfolgenden Grafik zeigt Zewo am Beispiel einer Schuldenberatungsstelle, ab welchem Schritt die Wirkungsmessung ansetzt. Wie bei Univation beginnt diese ab den Outcomes (s. Abb. 12)



Abbildung 12: Stufen der Wirkungsmessung anhand eines Beispiels (ebd.)

Wirkungsmodell nach Phineo

Kurz und Kubek von der Phineo gAG haben ein Kursbuch über die Implementierung einer Wirkungsanalyse für NPOs erstellt. Dieses Kursbuch enthält, wie das Logical Framework Approach, einen Problem- und einen Lösungsbaum. Es gibt einen Steuerungskreis mit drei Teilbereichen, wie bei der Stiftung Zewo, und es werden Bedarfs-, Umfeld- und Stakeholder*innenanalysen durchgeführt. Ziel des Ansatzes ist eine kontinuierliche Verbesserung der Arbeit. Auf die einzelnen Bestandteile der vollständigen Analyse wird in Kapitel 4.2.3 eingegangen. Hier wird die Wirkungstreppe von Phineo vorgestellt (s. Abb. 13). Sie ist ähnlich derer von Ottmann und König und Univation. Auch die Phineo gAG spricht ab der Stufe vier, die erste Stufe der Outcomes, von Wirkung. (Kurz und Kubek 2021, S. 5)

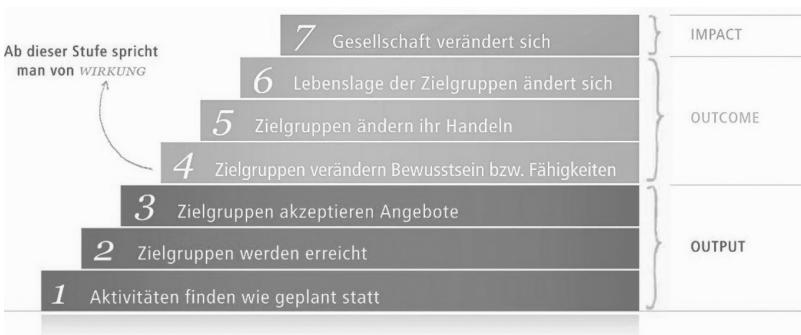


Abbildung 13: Wirkungstreppe von Phineo (ebd.)

Wirkungsmodell der New Economic Foundation (Nef)

Das Wirkungsmodell der Nef legt den Fokus auf die rein monetäre Darstellung der Wirkung. Es dient dazu, den SROI zu berechnen. Die Nef nutzt das Wirkmodell und den SROI zur strategischen Planung bzw. Entscheidungsfindung und zur Weiterentwicklung der Angebote. Besonders die Ergebnisse des SROI sollen zur Kommunikation über Wirkung genutzt werden und um Geldgeber*innen anzuwerben.

Die Berechnung des SROI erfolgt in sechs Schritten:

1. Anwendungsbereich festlegen und Stakeholder*innen identifizieren (z. B. durch eine Impact Map)
2. Abbildung der Outcomes und Darlegung des Zusammenhangs zwischen Input, Output und Outcome
3. Nachweis der Outcomes und Bewertung dieser
4. Feststellung von Wirkungen (Impact)
5. Berechnung des SROI
6. Berichterstattung, Nutzung und Einbettung. (übersetzt nach Nicholls et al. 2009, S. 9–10)

Wirkmodell nach Halfar und International Group of Controlling (IGC)

Halfar und die IGC stammen ursprünglich aus dem Profit-Bereich. Sie sind der Meinung, dass ihre Ansätze ebenfalls auf den NPO-Bereich in angepasster Weise übertragen werden können. Um die Wirkungsmessung vorzubereiten, schlagen sie zunächst die Nutzung eines Spinnenmodells vor (s. Abb. 14). Dieses dient dazu, die Organisation und ihre Ziele einzuordnen.

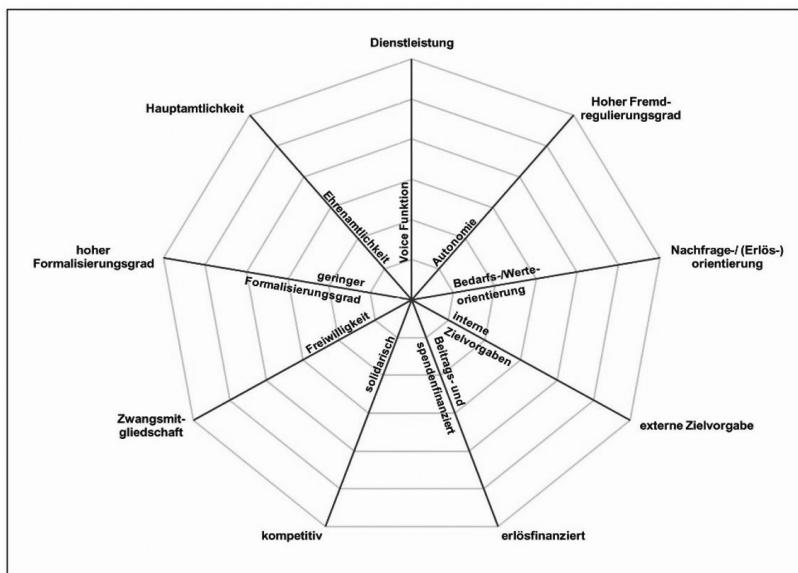


Abbildung 14: Grundmodell NPO-Spinne (Halfar und IGC-AG 2010, S. 17)

Mithilfe dieses Modells wird die IST- und die SOLL-Ausrichtung der NPO ermittelt. Anhand der Differenz zwischen den Ergebnissen können Ziele bestimmt werden. Im Anschluss daran werden die Stakeholder*innen und deren Erwartungen ermittelt. In der nachfolgenden Tabelle (Abb. 15) werden die verschiedenen Ziele pro Stakeholder*in in Auswirkungsstufen eingetragen.

	Direkte Wirkungs-empfänger Primary Customer	Weitere externe Stakeholder	Finanziers	Mitglieder/interne Stakeholder
Outcome				
Impact				
Effect				
Output				

Abbildung 15: NPO – Controllingsystematik: Wirkungsdimensionen mit Stakeholdern (ebd., S. 31)

Der untenstehenden Grafik (Abb. 16) ist zu entnehmen, ab wann Halfar und die IGC Wirkungscontrolling definieren. Hierbei ist anzumerken, dass die IGC die Begriffe Impact und Outcome anders definiert (s. Kapitel 3.1) als die vorher genannten Autor*innen.

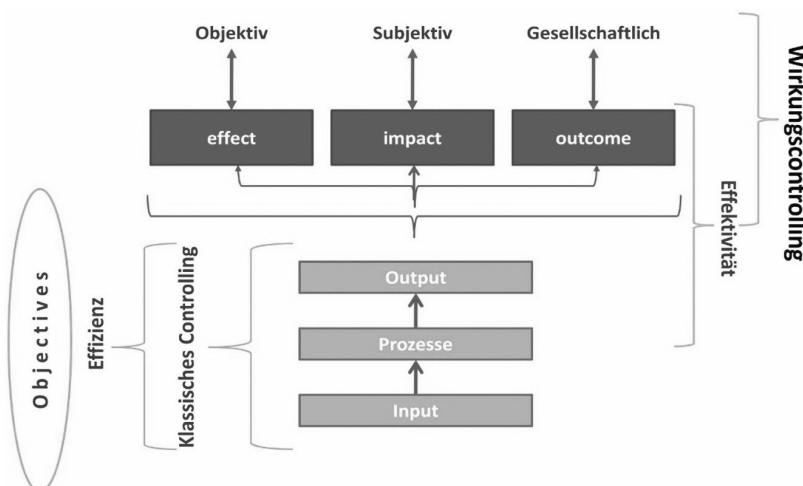


Abbildung 16: NPO-Wirkungsmodell (Halfar und IGC-AG 2010, S. 29)

Wirkungsmodell nach Grünhaus/Rauscher

Grünhaus und Rauscher beschreiben in einem umfassenden Working Paper verschiedene Ansätze. Ihre Theorie beruht darauf, dass die Erreichung von Wirkungen an der individuellen (z. B. Bedürfnispyramide, Maslow) und kollektiven (Menschenrechte, Sustainable Development Goals der UN) Bedürfnisbefriedigung orientiert ist.

Bestandteile ihres Wirkungsmodells sind die Wirkungskette (Abb. 17) und die Wirkungsbox (Abb. 18). Die Wirkungskette besteht, ähnlich wie bei bereits genannten Modellen, aus den IOOI-Komponenten. Zusätzlich werden die Aktivitäten und das Deadweight explizit genannt. Sie reihen mehrere Wirkungsketten für verschiedene zusammenhängende Projekte nebeneinander. Jede Wirkungskette ist zusätzlich für die Stakeholder*innen differenziert.

Die Wirkungsbox dient zur Informationssammlung und besteht aus mehreren Komponenten. Es gibt sechs Dimensionen (psychologisch, ökologisch, kulturell, politisch, sozial, ökonomisch), drei Ebenen (Mikro-, Meso-, Makroebene) und drei Zeitperspektiven (kurz-, mittel-, langfristig). (ebd., 20f)

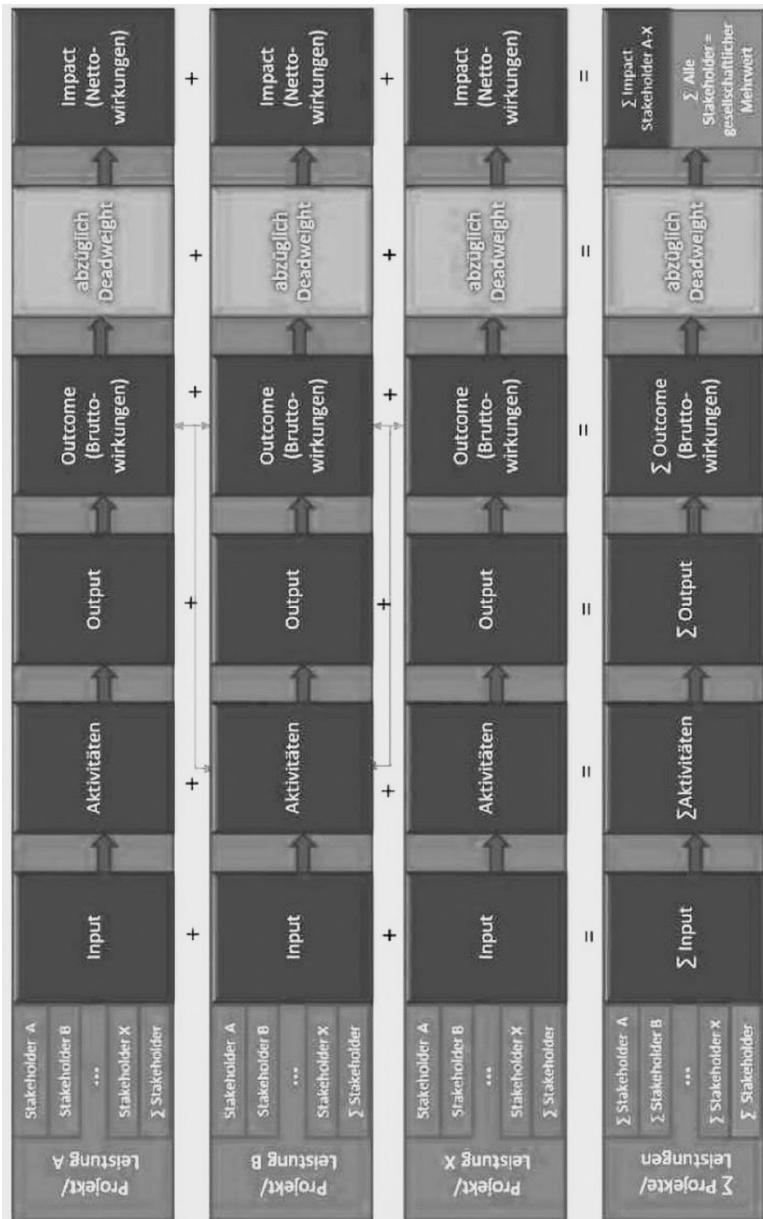
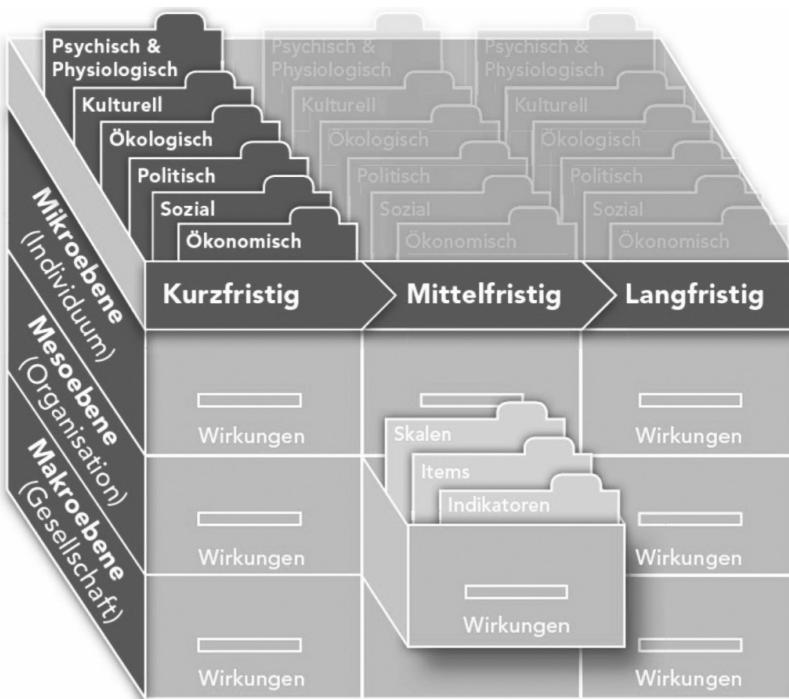


Abbildung 17: Einfaches Wirkungsmodell inklusive Stakeholder*innendifferenzierung (Grünhaus und Rauscher 2021, S. 21)



**IMPACT (NETTOWIRKUNG)=
OUTCOME (BRUTTOWIRKUNG) – DEADWEIGHT**

Abbildung 18: Wirkungsbox – Ebenen der Wirkungsbetrachtung (ebd., S. 24)

Wirkungsmodell von Ottmann und König

Die Vorgehensweise von Ottmann und König wurde bereits in Kapitel 3.3 beschrieben. Diese Grundelemente finden sich in Kapitel 4.2 erneut wieder, wenn die Wirkungsanalyse nach Ottmann und König beschrieben wird. Daher wird an dieser Stelle auf eine ausführliche Beschreibung verzichtet.

Auch wenn die Modelle sehr differieren, gibt es einige Schritte, die für alle Wirkungsmodelle bei der Entwicklung zu durchlaufen sind. Diese sind folgende:

1. Identifikation der zu messenden Maßnahmen
2. Identifikation aller Stakeholder*innen
3. Benennung der Wirkungsziele für die Zielgruppe (Outcome)
4. Benennung der Outputs, Aktivitäten und Inputs
5. Benennung von Income und Kontext
6. Sammlung von Stichpunkten für die einzelnen (Wirkungs-)Schritte
7. Präzisierung der Stichpunkte
8. Überprüfung der Zuordnung und Vollständigkeit
9. Darstellung der Wirkungszusammenhänge
10. Entwicklung von Indikatoren (ggfs. auch erst später bei der Wirkungsanalyse)
11. Regelmäßige Evaluierung/Anpassung des Modells (Baumgartner und Haunberger 2023, S. 119)

Je nachdem, für was das Wirkmodell genutzt werden soll und ob es in eine Wirkungsanalyse eingebettet wird, ist abhängig, wann das Modell erstellt wird. Soll es z. B. zur Programmentwicklung dienen, wird es im Vorfeld erstellt. Läuft das Projekt bereits, kann es auch für die Umsetzungsphase genutzt werden. Als Evaluationsinstrument kann es nach Projektende zum Einsatz kommen. (ebd., S. 118)

Wirkungsmodelle sind jeweils Grundlage für Wirkungsanalysen (Kapitel 4.1.2) und für wirkungsorientierte Steuerungsmodelle (Kapitel 4.2).

4.2.3 Wirkungsanalysen

In der Literatur ist die Differenzierung zwischen Wirkungsanalysen und wirkungsorientierten Steuerungsmodellen fließend. Dennoch wird hier versucht, eine Unterscheidung durchzuführen, da Wirkungsanalysen ein Baustein von Managementkonzepten sind, wohingegen wirkungsorientierte Steuerung eine „Haltung“ repräsentiert bzw. als eigenes Managementkonzept eingestuft werden kann.

Im Bereich der Wirkungsanalysen gibt es, wie bei den Wirkmodellen, unterschiedliche Ansätze. Die bekanntesten werden im Folgenden dargestellt.

Wirkungsanalyse nach Grünhaus/Rauscher

Grünhaus und Rauscher haben ein mehrschrittiges Modell (s. Abb. 19) entwickelt. Die Bestimmung des Zweckes der Analyse legt den Grundstein. Anschließend werden die Stakeholder*innen ermittelt und ein hypothetisches Wirkungsmodell erstellt. Darauf aufbauend werden die Wirkungen identifiziert und detailliert beschrieben, bevor die Messung der Breite und Intensität der Wirkungen erfolgt. Die ermittelten Wirkungen werden bewertet und nach einer Verdichtung dargestellt. Aus den Ergebnissen wird eine Strategie entwickelt, welche zusammen mit den Wirkungen kommuniziert wird. Diese Strategie dient anschließend zur Steuerung der Organisation. Die Schritte erweitern die im vorherigen Kapitel beschriebene Vorgehensweise von Baumgartner und Haunberger zur Erstellung eines Wirkmodells. Die Schritte drei bis sieben sind dabei optional umzusetzen.



Abbildung 19: Schritte einer Wirkungsanalyse eingebettet im Rahmen wichtiger Zwecke (Grünhaus und Rauscher 2021, S. 18)

Wirkungsanalyse von Ottmann und König

Ottmann und König beginnen die Wirkungsanalyse ebenfalls mit der Frage, welche Wirkungen gemessen werden sollen. Dabei weisen sie auf die Unterscheidung zwischen intendierten und nicht-intendierten Wirkungen hin. Der Fokus sollte auf den intendierten Wirkungen liegen, wobei die nicht-intendierten immer mit im Blick gehalten werden sollten. Nach dieser Festlegung wird eine Vorgehensweise für die Datenerhebung entwickelt. Anschließend werden Indikatoren gebildet, die zur Messbarkeit der Faktoren dienen. Danach werden Kriterien entwickelt, mittels welcher beurteilt werden kann, ob und in welchem Ausmaß die Maßnahmen wirksam sind oder nicht. (Ottmann und König 2023, S. 100) Als Tool für die Umsetzung einer Wirkungsanalyse haben sie den Wirkungsradar entwickelt (s. Abb. 20). Der Wirkungsradar gibt keine neuen Ansätze vor, sondern greift auf bestehende Black- und Whitebox-Verfahren zurück. Somit werden beide Grundfragen von Wirkungsmessung, ob und wie Wirkungen zustande kommen, beantwortet. Es handelt sich dabei um eine Empfehlung für ein stufenweises Vorgehen, wobei situativ entschieden werden soll, ob und inwieweit jede Stufe durchlaufen wird. Im Fokus dieser Vorgehensweise steht die Expertise der Fachkräfte und die Ergebnisse sollen zur fachlichen Weiterentwicklung dienen. Der Wirkungsradar gilt als ein Verfahren der realistischen Evaluation. Somit ist auch der Einbezug des Kontextes, wie oben bereits beschrieben, elementar. (ebd., S. 162)

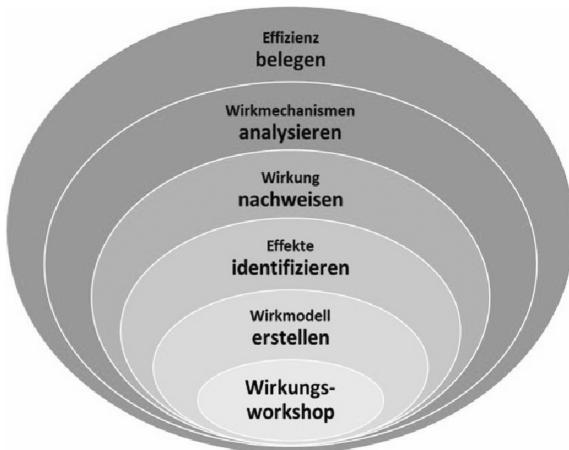


Abbildung 20: Der Wirkungsradar als empirische Toolbox für Wirkungsanalysen (Ottmann und König 2019, S. 68)

Als Grundlage für die Wirkungsforschung wird als Auftakt ein Workshop mit Mitarbeitenden und Führungskräften gesehen. Dieser dient als Basis zur Erstellung eines Wirkmodells (s. Kapitel 4.2.2). Anschließend werden die Effekte identifiziert. Dazu wird ein wirkungsorientiertes Monitoring empfohlen. Im nächsten Schritt werden Wirkungen nachgewiesen, z. B. durch quasi-experimentelle Designs. Sollte das nicht möglich sein, kann eine Wirkungsplausibilisierung durchgeführt werden. Anschließend werden die Wirkmechanismen analysiert, womit die Frage, wie die Maßnahme wirkt, beantwortet wird. Zum Abschluss kann die Effizienz, z. B. mittels SROI, belegt werden. (Ottmann und König 2023, 163ff)

Vorgehensmodell zur Wirkungsmessung der Stiftung ZEWO

Die Stiftung Zewo hat einen Kreislauf aus sechs Schritten für die Wirkungsmessung entwickelt. Der erste Schritt ist die Zieldefinition. I. d. R. wird diese bei einem Workshop unter Beteiligung der interessierten Parteien festgelegt. Schritt zwei ist die Entwicklung des Wirkungsmodells. Anschließend wird, in einem dritten Schritt, die Wirkungsmessung geplant. In Schritt vier werden Daten erhoben. Beim fünften Schritt wird die Wirkung beurteilt, bevor in Schritt sechs die Erkenntnisse genutzt werden. Anschließend werden diese wieder für die nächste Zieldefinition eingesetzt. Diese Schritte teilen sich auf die drei Phasen Planung, Umsetzung und Evaluation auf. (s. Abb. 21)

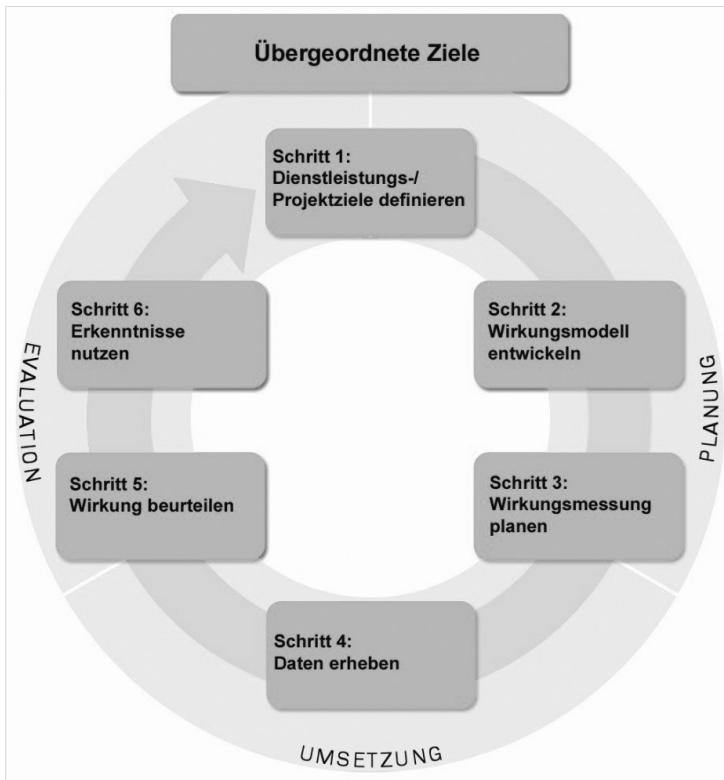


Abbildung 21: Wirkungsmessung anhand des Projektmanagement-Zyklus (Stiftung Zewo 2013)

Vorgehensmodell zur Wirkungsmessung nach Phineo

Kurz und Kubek haben ein „Logbuch“ für die Wirkungsmessung entwickelt. Es ist sehr praxisnah geschrieben. Ähnlich wie die Stiftung Zewo haben sie den Prozess in drei Hauptteile geteilt: Wirkung planen, Wirkung analysieren und Wirkung verbessern. Sie haben die einzelnen Phasen kleinschrittiger beschrieben. Im ersten Teil führen sie eine Bedarfs- und Umfeldanalyse sowie eine Stakeholder*innenanalyse durch. Diese Erkenntnisse fließen in einen Problembaum ein. Dieser dient als Grundlage zur Erstellung des Lösungsbaums, mit dem Ziele gesetzt werden, welche anschließend mit Hilfe der Wirkungstreppe (s. Kapitel 4.2.2) erarbeitet werden. In Teil zwei wird die

Wirkungsanalyse zunächst durch Festlegen der Rahmenbedingungen (wer führt sie durch?, welche Methoden werden genutzt?) und Entwicklung der Indikatoren vorbereitet. Dann erfolgt die Datenerhebung mittels quantitativer und qualitativer Methoden und die anschließende Auswertung. Im dritten Teil werden die Ergebnisse für die Weiterentwicklung genutzt und es erfolgt eine Veröffentlichung der Ergebnisse. (Kurz und Kubek 2021, 11ff)

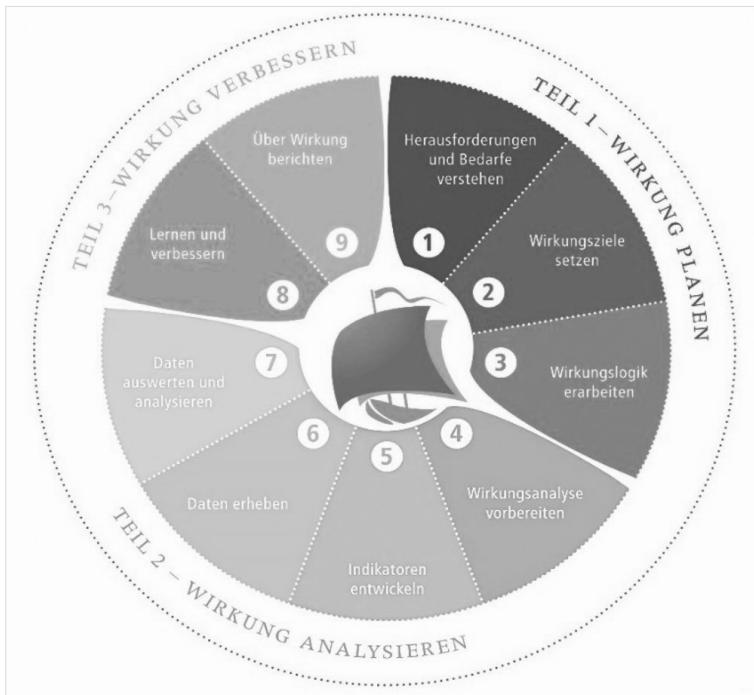


Abbildung 22: Wirkungsorientierung im Steuerungskreislauf:
Die Schritte der wirkungsorientierten Steuerung (ebd., S. 7)

Die einzelnen Bestandteile der Wirkungsanalysen differieren in geringem Ausmaß. Alle Autor*innen weisen zudem darauf hin, dass die Schritte als Richtlinie dienen und an die Gegebenheiten in der Praxis angepasst werden müssen. Alle haben gemeinsam, dass Indikatoren zur Messung entwickelt werden müssen. Darauf wird in Kapitel 4.2.5 näher eingegangen. Im folgenden Abschnitt werden die Stakeholder*innenperspektiven dargestellt.

4.2.4 Stakeholder*innenperspektiven

Um ein möglichst umfassendes Bild über wesentliche Wirkungen zu erhalten, ist es von hoher Bedeutung, möglichst eine große Bandbreite an Stakeholder*innen in den Prozess der Wirkungsanalyse einzubinden. I.d.R. besteht zu Beginn einer Analyse noch keine Klarheit darüber, wer zu der Gruppe der relevanten interessierten Parteien gehört. Daher empfiehlt es sich, zunächst eine größere Gruppe zu involvieren, auch wenn dieses zusätzlichen Aufwand bedeutet. (Grünhaus und Rauscher 2022, S. 515)

Wie schon bei den vorangegangenen Kapiteln ist auch bei der Stakeholder*innenanalyse grundlegend zu beachten, wer die Entscheidungen trifft. Die Indikatoren werden aus den verschiedenen Perspektiven unterschiedlich definiert. Darauf wird in Kapitel 4.2.5 eingegangen. Gleiches gilt für die Auswahl und Wichtung der Stakeholder*innen, welches abhängig vom Ziel der Forschung ist. Die Entscheidung, wer an der Wirkungsmessung mitwirkt, sollte optimalerweise bewusst von der Organisation entschieden und nicht von außen auferlegt werden. Ist z. B. das Ziel, die Arbeit bei Geldgeber*innen und Politik zu legitimieren, sollte der Fokus auf den Interessen der Leistungsträger*innen liegen. Geht es z. B. darum, einen Schwerpunkt der Sozialen Arbeit wie die Partizipation in den Mittelpunkt zu rücken, sind die Adressat*innen der Angebote die Hauptstakeholder*innen.

Die Durchführung der Stakeholder*innenanalyse sollte im Zusammenspiel von Führungsebene und Mitarbeitendenebene erfolgen. Es gibt zahlreiche Instrumente zur Erstellung einer Übersicht. In der nächsten Abbildung findet sich eine Tabelle über die relevanten Kategorien, die abgedeckt werden sollten. Es werden interne und externe Stakeholder*innen aufgelistet. Begonnen wird mit den Leistungsträger*innen, nachfolgend werden die Leitungskräfte und Mitarbeitenden der Organisation genannt. Anschließend werden die Zielgruppe und Personen in deren Umfeld aufgezählt. Abschließend wird die Fachöffentlichkeit, die allgemeine Öffentlichkeit sowie die Finanzierer*innen der Evaluation aufgelistet. Auf die Reihenfolge der Auflistung wird in Kapitel 6.2 Diskussion der Ergebnisse eingegangen.

Finanzierende Trägerschaft / Finanzierende der Evaluation	Finanzierende bzw. die Trägerschaft des zu evaluierenden Programms, allenfalls deckungsgleich mit der Instanz, welche die Kosten für die Evaluation trägt.
Verantwortliche bzw. Leitende	Eine zentrale Gruppe umfasst jene Personen, die für das evaluierte Programm bzw. für die untersuchte Intervention Verantwortung tragen.
Mitarbeitende, Vollzugsakteurinnen und -akteure	Die Personen, welche für die Umsetzung des untersuchten Evaluationsgegenstandes zuständig sind, gehören üblicherweise zu den Adressatinnen und Adressaten von Evaluationsresultaten.
Mitglieder der Zielgruppe	Die Mitglieder der Zielgruppe, welche von einem Programm oder einer Intervention erreicht werden, sind klassischerweise die «Programm-Betroffenen» (Balzer und Beywl 2018, S. 50).
Personen im Umfeld der Nutzenden	Im Umfeld der Adressatinnen und Adressaten lassen sich verschiedene Personen benennen, welche direkt oder indirekt vom untersuchten Programm oder einer Intervention betroffen sind.
Fachöffentlichkeit	Die Fachöffentlichkeit kann ein Interesse an Erkenntnissen zu einem Programm oder einer Intervention aufweisen.
Allgemeine Öffentlichkeit	Die allgemeine Öffentlichkeit ist möglicherweise in der Rolle als Steuerzahrende von einem Programm oder einer Intervention sowie durch das zu bearbeitende gesellschaftliche Problem betroffen.
Finanzierende der Evaluation	Möglicherweise wird die Evaluation von einer Akteurin oder einem Akteur finanziert, die nicht zugleich auch die Kosten des evaluierten Programms trägt.

Abbildung 23: Verschiedene Arten von Stakeholder*innen eines Programms (Baumgartner und Haunberger 2023, S. 64)

Wichtige Bestandteile der Analyse sind, neben der Ermittlung der relevanten internen und externen interessierten Parteien, zu bewerten, in welchem Bezug sie zur Organisation stehen und welches Interesse sie an der Maßnahme haben. Darüber hinaus ist einzuschätzen, wie stark ihr Einfluss möglicherweise ist. Auch sollte überlegt werden, aus welchem Grund sie einbezogen werden sollten. Das können z. B. folgende Gründe sein: Unterstützung bei der Umsetzung, Transparenz und Partizipation, Einbringen von normativen Perspektiven, methodische Gründe etc. (Baumgartner und Haunberger 2023, S. 63–66) Eine Dokumentation dieser Entscheidungen kann mithilfe der in Kapitel 4.2 auf Seite 51 vorgestellten Tabelle erfolgen. In Kapitel 5 befindet sich eine Übersicht über einige Stakeholder*innen des Arbeitsfeldes. Im nächsten Kapitel wird die Indikatorenbildung beschrieben.

4.2.5 Indikatorenbildung

Indikatoren und Kennzahlen dienen verschiedenen Zwecken. Sie können zur Informationsvermittlung und Kommunikationserleichterung beitragen. Außerdem haben sie Motivations- und Überprüfungsfunktionen. Sie können zur Planung und Kontrolle genutzt werden und dienen den Entscheidungsträgern bei der Organisationssteuerung. (Gruber 2018, S. 20) In der Literatur wird teilweise zwischen Kennzahlen und Indikatoren unterschieden, manchmal werden sie synonym verwendet. Die Autor*innen, die verschiedene Definitionen verwenden, begründen das u. a. dadurch, dass es sich bei Kennzahlen um verdichtete Maßgrößen aus absoluten und relativen Zahlen handelt, die aufbereitete quantifizierbare Informationen umfassen. Die Indikatoren gelten als Ersatzgröße, da sie einen nicht so hohen Grad an Reliabilität und Validität aufweisen, aber dafür leichter messbar sind. (ebd.; Kränzl-Nagl et al. 2019, 49f) Im Folgenden wird der Begriff Indikator, wie hier beschrieben, verwendet.

Die Bildung von Indikatoren ist ein elementarer Bestandteil der Wirkungsmessung. Sie geben die Zielrichtung der pädagogischen Maßnahmen vor. Je nach Festlegung des Zwecks der Wirkungsmessung und den gewünschten Zielsetzungen werden die Indikatoren unterschiedlich definiert. Wie oben erwähnt, spielt ebenfalls eine Rolle, wer die Kriterien festlegt. Im Optimalfall werden sie in einem Zusammenspiel von Führungskräften, pädagogischen Mitarbeitenden, Stakeholder*innen und Teilnehmenden definiert. Dieses ist bisher in der Praxis weitestgehend unüblich. Bspw. werden Ergebnisse aus Beteiligungsverfahren weniger stark genutzt als es sinnvoll wäre, es wird häufig über, aber nicht mit Personen geredet. (Polutta 2022, S. 230) Eine grundlegende Frage ist auch, was bereits als Wirkung gezählt wird und wer die Verantwortung für die Erreichung dieser trägt. (Albus und Ziegler 2013, 176f)

Bei der Entwicklung von Indikatoren sollte sich bewusst gemacht werden, dass eine Entscheidung für bestimmte Kriterien immer ein Ausschluss von anderen Kriterien bedeutet. Es handelt sich bei der Indikatorenentwicklung immer um eine positional objectivity, wenn eine große Anzahl von Personen im gleichen Kontext mit gleicher Informationsbasis ähnliche Kriterien entwickeln bzw. die Auswahl der Indikatoren akzeptieren würde. Dieses ist wertfrei zu betrachten. Gleichzeitig handelt es sich dabei

notwendigerweise um eine Auswahl von Wahrheiten und nicht um die einzige richtige Wahrheit, sondern um eine bewusste Entscheidung aufgrund von fachlichen und normativen Einschätzungen, die als relevant angesehen werden. (Albus und Ziegler 2013, S. 170) Eine Entscheidung z. B. für einen Indikator, der das kollektivistische Wohl misst (z. B. Kriminalitätsrate von Jugendlichen) kann dazu führen, dass das individuelle Wohl einer Person (z. B. Teilhabechance der Jugendlichen) zurückgestellt wird. Wenn der Fokus bspw. in erster Linie auf ökonomischen Aspekten liegt (Einsparung von SGB II Leistungen) bedeutet das nicht zwingend, dass die vermittelte arbeitslose Person eine sinnstiftende Arbeit gefunden hat, nur weil sie nicht mehr als arbeitssuchend geführt wird. (ebd., 170f) Deutlich wird die Diskussion auch z. B. mit Blick auf die Messung der Lebenszufriedenheit. Sie erfolgt aus einer strikt personenzentrierten Perspektive und somit rein subjektiv. Es wird dabei nicht hinterfragt, wie der Mensch seine Zufriedenheit definiert und was seine Bedürfnisse sind. Dabei werden u. a. zwei Schwächen der Forschung deutlich: das Unzufriedenheitsdilemma und das Zufriedenheitsparadoxon. (Boecker und Weber 2018, S. 12) Diese Aspekte sind besonders aus ethischer und machttheoretischer Perspektive zu beleuchten. In dieser Arbeit wird, wie in der Einleitung erwähnt, auf diese Betrachtungsweisen nicht näher eingegangen.

Die Entwicklung von Indikatoren ist u. a. aus dem zuvor genannten Grund anspruchsvoll und kritisch zu betrachten. Weitere Aspekte, die bei der Erstellung zu berücksichtigen sind, sind z. B., dass durch die Darstellung quantitativer Indikatoren eine Unterbewertung von qualitativen Wirkungen erfolgt. Außerdem darf der Arbeitsaufwand für die Erstellung der Indikatoren nicht höher sein als die Kernaufgaben der Tätigkeit, bzw. muss er in einem sinnvollen Verhältnis stehen. Auch die Schnelllebigkeit der Organisationen und der nötige Anpassungsbedarf sollten bedacht werden und wirken sich auf die Reflexion und Überarbeitung der Indikatoren aus. (Clifford 2015, 6f) Darüber hinaus ist das Ziel der Messung strategisch zu planen. Sollen die Indikatoren z. B. über die Wirkungsmessung hinaus auch zum Vergleich mit anderen Abteilungen/Institutionen genutzt werden, ist eine Skalierung der Indikatoren hilfreich. Dabei gilt zu beachten, dass genug Raum für konkrete und individuelle Anpassungen im Arbeitskontext bestehen bleiben (z. B. für situatives und fachli-

ches Agieren der Fachkräfte). (Schober und Rauscher 2014, S. 273) Es sollte darauf geachtet werden, dass die Vielfältigkeit der Organisationen und personenbezogenen Interventionen berücksichtigt werden und keine künstlich erzwungene Vergleichbarkeit erstellt wird, die für die eigene Organisation irrelevant ist (Clifford 2015, 6f).

Neben den Inhalten der Indikatoren und abhängig davon ist ebenfalls relevant, mit welchen Hilfsmitteln die Messungen erfolgen. Es wird zwischen quantitativen und qualitativen Erhebungsmethoden unterschieden. Je nach Wunsch bzgl. der Differenziertheit der Ergebnisse sind verschiedene Skalen (Nominal, Ordinal, Metrisch oder Rational) zu nutzen. (Ottmann und König 2023, S. 106)

Punkte, die bei der Erstellung von Indikatoren für eine Wirkungsanalyse wichtig sind, sind u. a., dass für jedes Wirkungsziel ein Indikator entwickelt wird, welcher sich klar von anderen unterscheidet. Für jedes Merkmal (Outcome oder Kontextfaktor) wird ein eigener Indikator benötigt. Diese Indikatoren müssen praxistauglich und messbar sein. Der Ressourcenbedarf für die Messung muss vorhanden und angemessen sein. Der Erhebungszeitraum muss realistisch und umsetzbar definiert werden. Der Indikator muss eine Veränderung darstellen und valide messen. Er muss darüber hinaus ständig und aktuell verfügbar sein. Außerdem sollte er so gestaltet sein, dass er nicht manipulierbar ist. (ebd., S. 115; Gruber 2018, S. 23)

In der Regel erfolgt die Indikatorenentwicklung durch Operationalisierung und ist somit deduktiv. Wie unter Punkt 4.2.2 beschrieben, werden im Prozess der Erstellung von Wirkmodellen zunächst die Impacts und Outcomes festlegt. Von diesen werden dann die Indikatoren abgeleitet. Dieser Vorgang erfolgt optimalerweise im Dialog aller Stakeholder*innen(-vertretungen). Ist der Impact bspw. die Reduzierung der Arbeitslosenquote und wäre der Outcome (in Abb. 24 „Gegenstand“) ein unterschriebener Arbeitsvertrag, dann wäre auf der Ebene der Dimension z. B. „gesundheitliche Vermittlungshemmnisse sind abgebaut“ anzusiedeln und ein Indikator könnte die Messung der (reduzierten) Krankheitstage sein.

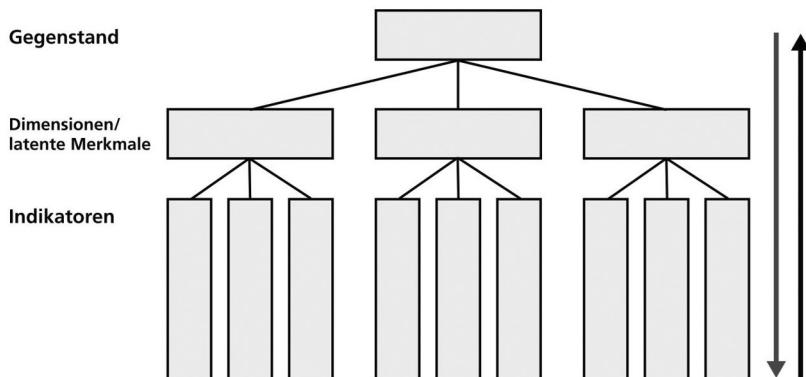


Abbildung 24: Prozess der Operationalisierung (Ottmann und König 2023, S. 105)

Um den Ressourcenaufwand möglichst gering zu halten, schlagen die Autoren des CSI-Gutachtens für die Träger der Freien Wohlfahrtspflege zugeschnittene Basis-Indikatoren vor. Diese würden bei dem Hauptträger entwickelt und auf die Ortsebene übertragen. Dadurch könnten aufwändige Validierungsstudien auf der Dachverbandsebene gebündelt werden. (Kehl et al. 2016, S. 63) Fraglich ist, wie praxis- und zeitnah diese Indikatoren zur Verfügung gestellt werden könnten.

4.3 Wirkungsorientierte Steuerungsmodelle

Wie in Kapitel 4.1 bereits erwähnt, gibt es unterschiedliche Gründe, warum Wirkungsanalysen durchgeführt werden. Unabhängig vom Anlass und dem damit verbundenen Aufwand, um solche Analysen durchzuführen, empfiehlt es sich, Wirkungsanalysen mit dem strategischen Management zu verbinden und in Form von wirkungsorientierten Steuerungsmodellen bzw. wirkungsorientiertem Management zu nutzen. (Grünhaus und Rauscher 2022, S. 507) Das Potenzial der Wirkungsorientierung kann im Rahmen eines Managementansatzes besser zur Missionserfüllung dienen als z. B. die reine Outputmessung (ebd., S. 508). Die Wirkungen werden dann zum jeweiligen Kern, um Entscheidungen zu treffen. (Schröder 2010, S. 103) Darüber hinaus dienen die Erkenntnisse der Wirkungsmessung

dazu, die Strategie der Organisation regelmäßig zu überprüfen und ggf. zu korrigieren. Durch eine dauerhafte Wirkungsmessung wird für das Arbeitsfeld eine umfängliche Analyse erstellt, die z. B. auch für Zukunftsprognosen und Entscheidungen über (Neu-)Ausrichtungen genutzt werden kann. Die fachliche Qualität der Arbeit sowie der Austausch darüber werden durch eine wirkungsorientierte Haltung verbessert. (Ottmann und König 2023, 38f)

Für die wirkungsorientierten Steuerungsmodelle bedarf es eines Umdenkens von einer inputorientierten Grundhaltung zu einem Blick auf die Wirkung und damit den Outcome und den Impact (Kränzl-Nagl et al. 2019, S. 23; Prinz 2019, S. 292). Dieses bedeutet, dass der Blick auf das Ende der Wirkungskette gerichtet und analysiert wird, was die Maßnahmen bei den Adressat*innen bewirken sollen. Es werden Wirkungsziele formuliert und anschließend Rückschlüsse auf die benötigten Ressourcen gezogen (Prinz 2019, S. 292).

Diese „konvertierte“ Denkweise ist in der folgenden Grafik dargestellt.



Abbildung 25: Strategische Pyramide mit Wirkungsmodell und Impact Map (ebd., S. 295)

Bei dem wirkungsorientierten Führungsansatz erfolgt die Planung ausgehend von der Vision und der Mission, die einen gesellschaftlichen

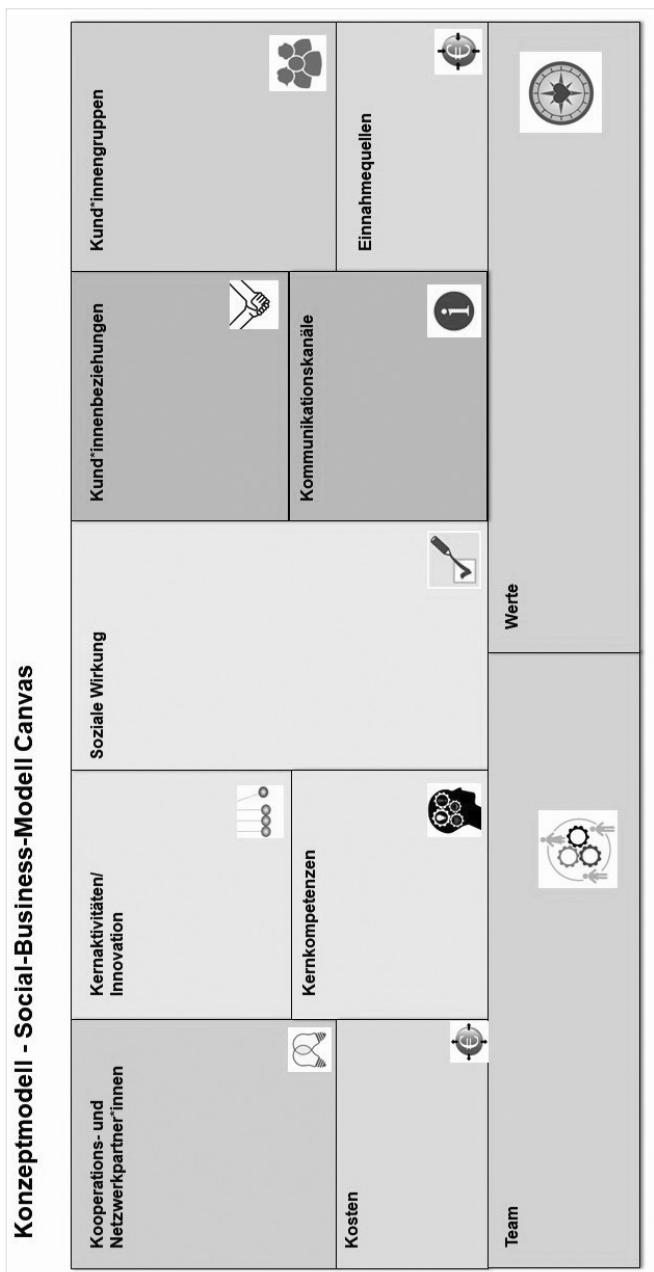
Mehrwert beschreiben sollten. Diese bilden die Grundlage für die strategischen (Wirkungs-)Ziele. Aus den strategischen Zielen werden operative Ziele und anschließend Maßnahmen abgeleitet. Daraus wird der Bedarf (Input) ermittelt. Die Steuerung in diesem Wirkungsmodell erfolgt anschließend in die entgegengesetzte Richtung. Im Optimalfall werden in die Planung die Interessen der Stakeholder*innen, welche Mitarbeitende, Adressat*innen, Leistungsträger*innen und weitere interessierte Parteien umfasst, mit einbezogen. Dieser Ansatz wirkt insofern dem machttheoretischen Ansatz entgegen, da die Ziele, Maßnahmen und Rahmenbedingungen aus der Mission abgeleitet sind. Kritisch ist zu betrachten, wer und für wen die Vision/Mission entwickelt hat/wurde.

Bei der Planung sind alle Wirkungen (intendierte, mögliche nicht-intendierte, positive wie negative) bestmöglich zu ermitteln, damit eine hohe Transparenz geschaffen werden kann und Entscheidungen dementsprechend getroffen werden können.

Dieser Ansatz läuft gegensätzlich zu der derzeit weitverbreiteten Vorgehensweise der Optimierung von Ressourcen, um möglichst viel Input zu haben, der viel Output generieren kann. (Prinz 2019, 293f)

Damit die Steuerung in der beschriebenen Art funktionieren kann, ist es wichtig, die Wirkungsmessung zu integrieren und nicht nur nebenherlaufen zu lassen. Es sollte sich um eine strategische Managementaufgabe handeln, die in sämtlichen Bereichen wie z. B. Controlling, Rechnungswesen, Mitarbeitendenführung etc. fest verankert ist. Diese Vorgehensweise geht über die reine Bildung von Indikatoren und Wirkungsketten hinaus. Erfolgreich umgesetzt führt sie zur Missionserfüllung und damit zur Verbesserung des gesellschaftlichen Mehrwerts. (Grünhaus und Rauscher 2022, S. 525)

Diese Herangehensweise sollte iterativ erfolgen und in regelmäßigen Abständen auch die Wertschaffungslogik der Organisation für die verschiedenen Stakeholder*innen in Frage stellen und weiterentwickeln. (Horak und Speckbacher 2022, S. 227) Für die regelmäßige Einschätzung der Wertschaffung und um Leistungsangebote zu bewerten kann das Social Business Model Canvas hilfreich sein. (ebd., S. 236–237) Es handelt sich dabei um ein aus dem Management abgewandeltes Instrument für die Soziale Arbeit, in dem übersichtlich dargestellt werden kann, auf



Stand: 28.12.2023

Abbildung 26: Social Business Modell Canvas, abgeleitet vom Business Model Canvas durch TUM des CV Kölns angepasst auf den LB BuA

was sich eine Organisation fokussieren will und welche Aufgaben und Themen nicht behandelt werden sollen. Das Innovationsteam (TIM) des Caritasverbandes für die Stadt Köln e. V. hat ein solches Modell für den Verband abgeleitet (s. Abb. 26).

Ein Ansatz für wirkungsorientierte Steuerung ist z. B. der des Design Thinkings. Dieses Instrument ist ebenfalls aus dem Profit-Bereich und wird dort für die Entwicklung von Lösungen für bestehende Herausforderungen genutzt. Es besteht aus verschiedenen Schritten:

„Verstehen, Untersuchen, Synthese, Ideenfindung, Prototyping, Testen“ (Ottmann und König 2023, S. 79). Ottmann und König haben die Phasen in der nachfolgenden Grafik angepasst dargestellt.

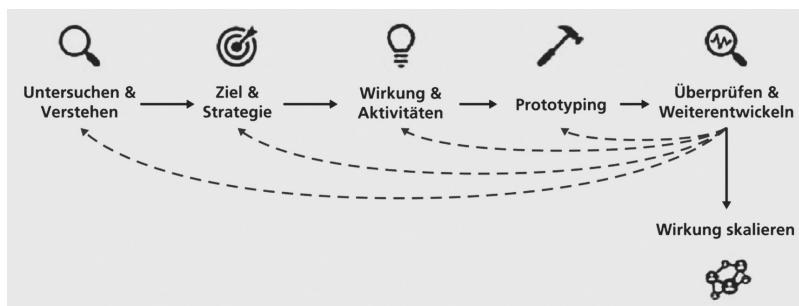


Abbildung 27: Prozess der wirkungsorientierten Konzeptionsentwicklung (ebd.)

Die Vorgehensweise zeichnet sich durch ständige Rückkopplungsprozesse aus, um das Ergebnis regelmäßig zu überprüfen und weiterzuentwickeln. Die Phasen müssen dabei nicht in der aufgelisteten Reihenfolge durchlaufen werden. (ebd., S. 80) Im Design Thinking kann auf verschiedene Methoden zurückgegriffen und mit anderen kombiniert werden, z. B. auf die Problem- und Lösungsbäume des Logical Framework Ansatzes (s. Kapitel 4.2.2 und 4.2.3).

4.4 Implementierung von Wirkungsmessung

Es handelt sich bei der Implementierung einer Wirkungsanalyse oder von wirkungsorientierter Steuerung um einen umfangreichen Organisationsentwicklungsprozess, der besonders in einer linien-hierarchischen Verbandsstruktur ausführlich kommuniziert und abgesprochen werden muss. Für diesen Ablauf ist es wichtig, alle Stakeholder*innenperspektiven mit einzubeziehen. Die Implementierung muss sorgfältig durchdacht und geplant werden. Es gibt dabei mehrere Gesichtspunkte, die zu berücksichtigen sind. Ein Blickwinkel ist der auf systemische Aspekte. Unter anderem sind dabei die Auswirkungen autopoetischer Systeme, der Drang nach Selbsterhaltung, zu berücksichtigen. (Wollnik 1998, 141f) Es ist davon auszugehen, dass Widerstände bei der Einführung entstehen werden. Dieses kann mittels Ansätzen aus den Implementierungswissenschaften aufgefangen werden. Einflussfaktoren auf fünf Ebenen sind in den Blick zu nehmen: involvierte Individuen (Fachkräfte, Klient*innen, andere Akteur*innen), welche bewusst oder unbewusst auf Neuerungen reagieren und den Prozess steuern; Merkmale der Personen und deren soziale Interaktion, damit sind bspw. gruppendiffusivische Prozesse oder Kommunikationsstrukturen gemeint; interner Kontext, „in dem sich die Individuen bewegen“ (z. B. Organisation, Unterkunft, Team oder z. B. Organisationskultur); externer Kontext (Umweltbedingungen, z. B. Konkurrenz, Gesetze); Merkmale der Neuerung inkl. veränderbarer und nicht veränderbarer Eigenschaften (z. B. Komplexität oder objektiver Nutzen) und der Implementationsprozess in sich selbst (z. B. Merkmale, Gestaltung und Planung des Prozesses). (Hoben 2015, 151)

Bei der Implementierung von Projekten bzw. von Organisationsprozessen ist es empfehlenswert, ausreichend Zeit für Diskussionsphasen einzuplanen, damit alle Mitarbeitenden Raum erhalten, sich einbringen zu können. Der Prozess sollte mittels einer kontingenten Projektführung umgesetzt werden, in der Rückschritte und Wiederholungsschleifen von Beginn an proaktiv eingeplant werden. Dies ermöglicht die Offenheit für mehrere Lösungswege, da davon auszugehen ist, dass es nicht die eine richtige Lösung gibt. (Kühl 2016, S. 37; Metternich 2023, S. 8) Da es sich bei Implementierungsprozessen um soziale Prozesse handelt, ist davon

auszugehen, dass der Verlauf eher „chaotisch“ statt kontinuierlich-linear verläuft und daher „nur begrenzt berechen-, plan- und steuerbar“ (Hoben 2015, S. 150) ist. Die Einflussfaktoren auf den Prozess sind i. d. R. multipel, sich verändernd und komplex interagierend. Dadurch ergeben sich regelmäßig neue Phänomene, die mit in den Blick genommen werden müssen. (ebd.) Eine zirkuläre Vorgehensweise mit regelmäßigen Evaluationszyklen wird empfohlen. Diese Zyklen bestehen z. B. u. a. aus den Schritten „Verständigen, Entwickeln, Analysieren, Weiterentwickeln, Berichten und Skalieren“ (Ottmann und König 2023, 30) oder, wie in Kapitel 4.3 beschrieben, aus den Schritten des Design Thinkings. Zusätzlich wird empfohlen, die Implementierung zunächst in einer Einrichtung zu erproben, bevor eine flächendeckende Umsetzung erfolgt. (Ottmann und König 2023, S. 33) Dieses hat mehrere Gründe. Erstens kann ein Abbruch leichter umgesetzt werden, wenn die Methodik nicht zielführend eingesetzt werden kann, zweitens kann eine Art Pilotversion entwickelt werden, die später leichter übertragen werden kann und drittens ist der Ressourcenaufwand zunächst für eine Abteilung geringer als für mehrere Bereiche gleichzeitig.

Für die Einführung von Wirkungsmessung gibt es darüber hinaus weitere Empfehlungen, die im Folgenden dargestellt werden. Zunächst sollte überlegt werden, wozu Wirkungsmessung implementiert werden soll und welche Erwartungen an das Ergebnis gestellt werden. Je nach Ziel macht es Sinn, alternative, ggfs. niedrigschwellige Ansätze zu prüfen. Als weiteren Schritt ist die Implementierungstiefe zu beurteilen. Unter Umständen reicht eine reine oberflächliche Darstellung von Wirkung oder die Diskussion über Wirkannahmen als Alternative zu einer vollständigen Wirkungsmessung, um in einen Austausch zu kommen. Anschließend sollte die Einbindung der interessierten Personen erfolgen. Dabei ist ebenfalls zu entscheiden, mit welchem Zweck, welchem Ausmaß und mit welcher Methode dieses erfolgen soll. Es ist außerdem abzuwegen, mit welchen Kosten bzw. Aufwand zu rechnen ist und welche Aufgaben stattdessen vernachlässigt werden können, sofern keine weiteren Ressourcen zur Verfügung stehen. Auch bei der Einführung von Wirkungsmessung wird die Anwendung von etablierten Projektmanagement- bzw. Implementierungsprozessen empfohlen. Dabei gilt es, negative bzw. unerwünschte Effekte zu beobachten und zu evaluieren und ggfs. nachzusteuern. Widerstän-

de sollten dabei zur Reflexion genutzt werden. Es empfiehlt sich darüber hinaus, den Prozess kontinuierlich mit regelmäßigen Schleifen zu implementieren, anstatt als adhoc-Projekt durchzuführen, welches nicht mehr nachverfolgt wird. (Gössler 2016, 130ff)

Sollten nicht ausreichend Ressourcen zur Verfügung stehen oder ist eine umfassende Wirkungsanalyse nicht notwendig, reicht häufig eine „systematische Auseinandersetzung mit der Wirkungsmessung“ (Kehl et al. 2018, S. 283) mittels eines Wirkmodells in Hinblick auf die Visionen und Strategien einer Organisation aus. Statt einer evidenzbasierten Forschung kann auf plausible Schätzungen zurückgegriffen werden. Diese Vorgehensweise trägt bereits entscheidend zur Legitimation der Maßnahmen bei und kann der strategischen Entscheidungsfindung dienen. Dabei kann auf die oben genannten Methoden zurückgegriffen werden, um in Projektteams Wirklogiken systematisch herzuleiten. (ebd.)

4.5 Folgen von und ein kritischer Blick auf Wirkungsorientierung

Die Implementierung von Wirkungsorientierung hat Folgen für die Arbeit, sofern sie erfolgreich umgesetzt und nicht vom System abgestoßen wurde. Auf der Input-Seite wird die Konsequenz sein, dass Leitungskräfte bzgl. wirkungsorientierter Steuerung Wissen aufbauen müssen. Zusätzlich bedarf es der Schulung von Mitarbeitenden mit Blick auf die Umsetzung der Wirkungsmessung. Da wirkungsorientiertes Steuern Auswirkungen auf die Haltung in der Organisation hat, ist eine Anpassung der Strukturen und Abläufe sowie die Nutzung anderer Methoden/Instrumente erforderlich. Darüber hinaus sind ergänzende oder eine Umstrukturierung der Ressourcen nötig, um Wirkungsmessung umzusetzen.

Es wird davon ausgegangen, dass sich Leistungsverträge zu Wirkungsverträgen entwickeln werden. Hierbei kann diskutiert werden, ob die Änderung des Managements zu anderen Verträgen führen wird oder der Bedarf der Anpassung des Managements an die neuen Verträge besteht. (Schober und Rauscher 2014, S. 275)

Die positiven Auswirkungen von Wirkungsmessung sind eine effektivere und effizientere Nutzung von Ressourcen sowie zielgerichtete Maß-

nahmen für Adressat*innen, die sich aufgrund der Wirkungsergebnisse für oder gegen Angebote entscheiden können (sofern eine Wahlmöglichkeit besteht). Hier kann kritisch hinterfragt werden, ob diese Resultate nicht auch anders erreicht werden könnten.

Mögliche negative Folgen können entstehen, wenn die Messungen als Selbstzweck in Machtstrukturen genutzt werden und die Zielerreichung eine Eigendynamik entwickelt. Zu einschränkende Wirkungsmessungen können darüber hinaus zu einer Rückkehr zu standardisierten Beratungsprozessen führen, ohne Flexibilität für den Einsatz von für die Situation angemessenem Fachwissen. (Freier 2023, S. 24) Dieses wird am Beispiel der Zielvorgaben des Jobcenters für Wiedereingliederungsmaßnahmen deutlich. Das Jobcenter hat zum Ziel, möglichst viele Menschen durch die Maßnahmen zu „schleusen“ und hohe Vermittlungsquoten zu erreichen. Dieses Tempo widerspricht häufig der pädagogischen Einschätzung der Fachkräfte und setzt die Teilnehmer*innen unter Druck.

Auf der politischen Ebene wird kritisiert, dass Wirkungsmessung zur „Verschleierung politischer Entscheidungen für Einsparmaßnahmen“ (Finis Siegler 2021, S. 39) dient. Positiv kann es als effektiver und effizienter Einsatz von Steuermitteln gesehen werden.

