

Wirkungen und Qualität sozialer Dienstleistungen

Das Modellvorhaben „Kein Kind zurücklassen“ will das Hammer Bildungs- und Sozialsystem zugunsten gelingender Entwicklung von Kindern und Jugendlichen effektiver gestalten. Mit den nachstehenden Überlegungen will der Fachbeirat einen fachlich-konzeptionellen Rahmen für die Auseinandersetzung um die Wirkungen und die Qualität sozialer Dienstleistungen und die damit verbundene Diskussion um die Evaluation von Maßnahmen und Projekten schaffen.

1 Anforderungen und Probleme

Jedes professionelle Handeln ist mit einer Wirkungserwartung verbunden; man will mit dem Handeln Ziele erreichen, einen Effekt erzeugen. Der Erfolg der Handlung wird daran gemessen, ob bzw. wie stark der erwartete Effekt eingetreten ist.

Somit erscheint es verständlich, dass für die Bewertung des beruflichen Handelns (Erfolg, Zielerreichung) die Frage nach der Wirksamkeit des Handelns von Bedeutung ist:

- Stimmen die tatsächlichen Wirkungen mit erhofften und geplanten Wirkungen überein?
- Sind neben den beabsichtigten Wirkungen positive oder problematische Nebenwirkungen aufgetreten?
- Haben kurzfristige Wirkungen auch mittelfristig Bestand?
- Sind die Wirkungen bei verschiedenen Adressaten ähnlich oder unterschiedlich?

Wirkung bzw. wirkungsorientierte Steuerung wird in der Sozialpolitik mit zunehmender Relevanz diskutiert. Es gibt eine veränderte Akzentsetzung für den gesellschaftlichen Umgang mit sozialer Arbeit: die nachzuweisende Wirksamkeit gilt zunehmend als Legitimationskategorie für soziale Einrichtungen.

Während in der Vergangenheit die Überprüfbarkeit sozialen Handelns z.T. vehement bestritten wurde, im Wesentlichen mit dem Argument der Unvergleichbarkeit von naturwissenschaftlichen Gegenständen und Prozessen in sozialen Systemen, verbreitert sich heute die grundsätzliche Anerkennung der Möglichkeit und Notwendigkeit der Überprüfung von Wirkungen sozialer Dienstleistungen. Die Debatte verlagert sich auf die Frage, wie dies sinnvoll und mit vertretbarem Aufwand möglich ist. Diese Debatte muss aufmerksam und kritisch verfolgt werden. Fehlentwicklungen sind aus Sicht des Fachbeirates:

- Die unmittelbare Verkoppelung von Wirkungsanalyse und finanzieller Steuerung: Sie ist insbesondere dann problematisch, wenn mit der Überprüfung der Wirksamkeit von Maßnahmen Einsparungsziele verbunden sind. Sie ist auch deshalb problematisch, weil kurzschlüssige Kopplungen (z.B. „keine Wirkungen – kein Geld“) einen starken verfälschenden Effekt auf jede Überprüfung haben.
- Die Wirkung in der Handlungslogik der Sozialen Arbeit ist ein zentraler Bezugspunkt. Allerdings ist die Diskussion in der Sozialen Arbeit um die Wirkungsorientierung von starken Ambivalenzen geprägt:
 - Zum Teil müssen naive, monokausale Vorstellungen über die Überprüfbarkeit von sozialen Dienstleistungen abgewehrt werden, die von der Öffentlichkeit, z.T. auch von Entscheidungsträgern und Politik an die Akteure herangetragen

werden. Entsprechende Ansätze gab und gibt es auch in Hamm. Die Herausforderung besteht darin, aussagekräftige und gleichzeitig mit vertretbarem Aufwand erhebbare Indikatoren zu finden.

- Es gibt eine Reihe typischer Fehler, die in Verbindung mit der Evaluation von Maßnahmen und einem ungeeigneten Wirkungsverständnis auftauchen und die durch das vorliegende Papier nicht begünstigt, sondern eher verhindert werden sollten. Dazu gehören zum Beispiel:
 - Ziele werden zu hochgesteckt, zum Beispiel um (vermeintliche) Erwartungen von Fördermittelgebern zu erfüllen.
 - Ziele werden nicht konkret genug formuliert.
 - Maßnahmen werden nicht individuell bezogen auf Ziele und Zielgruppen entwickelt, sondern „von der Stange“ genommen (gemäß dem Satz von Paul Watzlawik: „Wenn du einen Hammer hast, behandelst du jedes Problem als Nagel“.)
 - Es werden gut verfügbare Indikatoren ohne Aussagekraft definiert.
- Die Formulierung und die Erreichung von Zielen sind als Prozess zu betrachten, und dieser Prozess muss sich in der Evaluation abbilden. Insofern müssen Ziele auch im Prozess überprüft und verändert werden können.
- Projektüberprüfung und Projektevaluation müssen eingebettet sein in lernende Organisationen und Lernende Organisationsstrukturen. Die Neugierde und das Interesse an der gemeinsamen Weiterentwicklung sind ausschlaggebend, nicht der Versuch, Einsparungen zu erzielen.
- Eine gute Evaluation von sozialen Projekten erfordert einen hohen Aufwand. Dieser muss in der Zeit- und Ressourcenplanung berücksichtigt sein. Auf der anderen Seite können in einem systemischen Verständnis Evaluationsmethoden als Bestandteil sozialer Prozesse und Entwicklungen genutzt werden (zum Beispiel kann die Befragung von Kursteilnehmern nach dem, was sie gelernt haben, zu einer Vertiefung des Lerninhaltes führen).

In diesem Spannungsfeld entwickelt der Fachbeirat Vorschläge für die Fachdiskussion in Hamm.

2 Definitionen

2.1 „Personenbezogene soziale Dienstleistungen“

Der Begriff der „personenbezogenen sozialen Dienstleistungen“ gilt als Oberbegriff für die unterschiedlichen Aktivitäten und Arbeitsformen, die durch das öffentliche Sozial-, Bildungs- und Hilfesystem (durch freie und öffentliche Träger) erbracht werden. Eine „personenbezogene soziale Dienstleistung“ ist u.a. durch die folgenden Merkmale definiert:

- sie ist öffentlich verfasst
- sie wird beruflich oder ehrenamtlich erbracht
- und sie macht die Koproduktion der Konsumenten/innen erforderlich.

Im Rahmen der Umsetzung z.B. der Hilfen zur Erziehung gem. § 27 SGB VIII gibt es eine Verantwortungsgemeinschaft (Jugendamt, freier Träger und Adressat), die für den Erfolg der Leistung verantwortlich ist. Ein besonderes Merkmal der personenbezogenen sozialen Dienstleistung liegt in der zeitgleichen Produktion und Konsumtion, die alle nachfolgenden Prozesse beeinflusst.

2.2 Qualität

Allgemein wird die Qualität eines Produktes oder einer Dienstleistung (nach DIN EN ISO 9004) durch die Gesamtheit von Eigenschaften und Merkmalen, die sich auf die Erfüllung festgelegter oder vorausgesetzter Erfordernisse beziehen, bestimmt. Bei der Qualitätsbeurteilung einer Dienstleistung lassen sich drei Dimensionen unterscheiden:

- (1) Die Strukturqualität bezieht sich auf die Rahmenbedingungen bei der Erstellung einer Dienstleistung, z.B. die personelle Ausstattung, die Infrastruktur, die Erreichbarkeit u.am.
- (2) Die Prozessqualität beschreibt die Gestaltung der Abläufe zur Erbringung der Dienstleistung.
- (3) Die Ergebnisqualität betrachtet die erzielten Resultate und Effekte.

2.3 Wirkung

Die Erwartung, mit einer Handlung einen bestimmten Effekt zu erzeugen, bildet dementsprechend die Grundlage für die Beurteilung, ob eine Handlung erfolgreich war: Ist der erwartete Effekt eingetreten? Hat die Handlung „Wirkung“ erzeugt?

Die Wirksamkeit der sozialen Dienstleistung orientiert sich somit an vielfältigen Kontextbedingungen und Wirkfaktoren, so dass eine mono-kausale Bewertung einer „pädagogischen Intervention“ im Sinne einer einfachen Zustandsveränderung Vorher vs. Nachher für eine Wirkungsevaluation nicht ausreichend ist.

„Wirkungen sind intendierte Zustandsänderungen, die beobachtbar, beschreibbar und kommunizierbar sind und nach plausiblen und hypothesengeleiteten Annahmen über nachvollziehbare Zusammenhänge bewertet werden können“ (vgl. Wirkungsorientierte Jugendhilfe Band 9 , Seite 22).

2.4 Plausibilität

Die Grundlage einer jeden Evaluation bilden Annahmen über Wechselwirkungen. Diese Wechselwirkungen zwischen unterschiedlichen Faktoren sind in der Regel nicht als kausale Faktoren zu identifizieren. Diese Faktoren machen die Festlegung von Wirkungskategorien zu einem schwierigen, tastenden Vorgang. In der Regel wirken viele Faktoren auf die Lebenssituation und auf die helfende/intervenierende Handlungssituation der sozialen Arbeit ein, so dass kausale Faktoren für Wirkung nur schwer zu bestimmen sind:

- Soziale Arbeit es meist mit „wildem“, relativ wenig strukturierten Problemen zu tun hat
- „Wilde Probleme“ sich nicht definitiv und endgültig beschreiben lassen
- Sie sich häufig verändern und einer besonderen Dynamik unterliegen
- Es meist ungewiss ist, wann das „wilde Problem“ als gelöst angesehen werden kann
- Häufig vorläufige Zwischenlösungen erreicht werden

Wir sind aber aufgrund der o.g. Zielsetzungen auf eine empirische Basis zur besseren Einschätzung der Wirkungsintensität ihres Handelns angewiesen. Wir bewegen uns auf der Ebene von empirischen Hinweisen zu Plausibilitäten. Von den Effekten, die in einer Vorher – Nachher – Messung zutage treten, wird mit einem gewissen Grad von Wahrscheinlichkeit (Plausibilität) angenommen, dass sie in einem Zusammenhang mit der Intervention stehen.

Je differenzierter die Annahmen sind, und je besser einzelne Faktoren z.B. durch wissenschaftliche Studien als besonders hilfreich bzw. wirksam identifiziert sind, um so besser sind Aussagen über die Wirksamkeit sozialer Maßnahmen zu treffen.

2.5 Zufriedenheit und Wirkung

Die Zufriedenheit mit einer Leistung ist nicht identisch mit der Wirkung einer Leistung. Wenn Menschen mit den Personen, der Atmosphäre, etc. zufrieden sind, hat dies nicht automatisch zur Folge, dass sie im Anschluss an eine Intervention ihre Probleme besser bewältigen können und dabei auch Nachhaltigkeit herzustellen vermögen.

Dies wäre aber die gewünschte und erwartete Wirkung.

Soziale Dienstleistungen hängen substanziell von der Bereitschaft zur Koproduktion auf Seiten der Nutzer ab. Deshalb sollte sich die Nutzerperspektive in der Evaluation niederschlagen. Somit kann die Zufriedenheit einen wichtigen Baustein innerhalb der Bewertung eines Angebotes darstellen.

2.6 Zeitdimension

Jede Wirkungsmessung richtet sich auf bestimmte Zeitpunkte eines Programmverlaufes. Dies bedeutet, dass sich zu einem bestimmten Zeitpunkt eine bestimmte Wirkung zeigen soll. Zeiträume und Zeitpunkte einer Messung müssen begründet ausgewählt werden. Die Wirkung ist ein wichtiger Gegenstand von Evaluationen. Die genannten Schwierigkeiten sollen nicht dazu führen, sich auf Konzept-, Struktur-, und Prozessevaluationen zu beschränken.

Das Handlungsfeld der Sozialen Arbeit ist nicht technologisierbar und zeichnet sich durch eine Unsicherheit hinsichtlich der genauen Problemkonturen aus. Aus diesem Grund ist sie dem engen und traditionellen Verständnis von Evidenz nicht zugänglich.

Wirkungsevaluation in der sozialen Arbeit ist letztlich immer auf das Herstellen von Plausibilitäten angewiesen. Wirkungsevaluationen können helfen, diese Plausibilitäten zu schaffen oder sie zu erhöhen. Diese Option ist der Mühe wert!

2.7 Evaluation

„Evaluation“ bezeichnet die methodische Erfassung und Bewertung (sowie Dokumentation) von Maßnahmen, Projekten, Programmen, Institutionen, Organisationen, etc.

Die Evaluation vollzieht sich somit auf mehreren Ebenen:

- auf der Ebene von Auftraggeber und Auftragnehmer
- auf der Ebene der Adressaten
- auf der sozialräumlichen Ebene und
- auf der kommunalen Ebene

Für die Evaluation notwendige (Wirkungs-)Indikatoren müssen in diesem Kontext wohlüberlegt formuliert sein und beinhalten immer den Aspekt der Kundenzufriedenheit. Die Kundenzufriedenheit ist insofern reflexiv auf den Anteil des Adressaten und die Anteile der sozialen Dienstleister zu beziehen, da in den einzelnen Co-Produktionsprozessen Produktion und Konsumtion zeitgleich erfolgen und die folgenden Prozesse maßgeblich beeinflussen.

Für das Controlling bedeutet dies, dass Kennzahlen zur Wirkung vor allem globaler Natur sind und ihre Formulierung insbesondere langfristige, systemübergreifende

Zustandsveränderungen in den Focus nimmt. Eine Besonderheit in diesem Kontext ist die „nicht-lineare“ Wirkungsentfaltung der sozialen Dienstleistung.

Für die Strategien von Auftraggebern und Auftragnehmern zeigt sich die Formulierung von SROIs – Social Return On Invest – als hilfreich.

2.8 Controlling

„Der Begriff stammt aus der Betriebswirtschaft und bezeichnet eine Managementfunktion, deren Aufgabe die Koordination von Planungs-, Entscheidungs-, Durchsetzungs- und Kontrollprozessen eines Unternehmens ist (Gerlich, 1999; Pech, 2001). Controlling kann sich auf verschiedene Bereiche beziehen, wie etwa Kosten, Personal oder Bildung (Bildungscontrolling), und betrachtet diese als Steuerungssysteme, die mit Hilfe von Instrumenten wie Kennzahlensystemen, Betriebsrechnungen aber auch Evaluationen rational regelbar sind. Der Begriff wird heute nicht mehr primär nur auf monetäre Indikatoren bezogen. (...) Im Gegensatz zum Monitoring hat das Controlling eine aktivere Rolle. Daten werden nicht nur gesammelt, sondern auch interpretiert und explizit für Steuerungsprozesse genutzt.“ (<http://www.evaluation.de/glossary#c>).

2.9 Monitoring

„Routinemäßige, kontinuierliche und systematische Sammlung von Vergleichsdaten mit dem Zweck, Trends und Veränderungen sichtbar zu machen.

Die Datensammlung beim Monitoring ist nicht willkürlich, sondern beschränkt sich auf bestimmte Indikatoren, die regelmäßig erfasst werden. Im Gegensatz zur Evaluation nimmt Monitoring keine Bewertung von erhobenen Daten vor, hat also eine rein beobachtende Aufgabe. So verstanden würde es sich etwa bei den PISA-Studien i.e.S. eher um Monitoring als um Evaluation handeln“ (<http://www.evaluation.de/glossary#c>).

3 Evaluationsdesigns für Wirkungsevaluationen

Um soziale Projekte und Bildungsmaßnahmen zu evaluieren, gibt es unterschiedliche Ansätze. Das zuvor Gesagte macht klar, dass in jedem sozialen Prozess komplexe Wechselwirkungen stattfinden. Ziel einer Evaluation muss es sein, möglichst genau die Wirkungen zielgerichteter Interventionen zu identifizieren, ohne allerdings einem mechanistischen oder monokausalen Wirkungsverständnis aufzusitzen.

3.1 Das randomisierte Experiment

- Gruppe, Kontrollgruppe oder Vergleichsgruppe
- Vorher-Nachher - Messung und Vergleich der Gruppen
- nur unter kontrollierten Bedingungen
- in der Sozialen Arbeit nicht umsetzbar

3.2 Das quasi-experimentelle Verfahren

- Messung im Gruppenvergleich
- Gruppenbildung eher zufällig
- keine kontrollierten Bedingungen

„Je praxisnäher Evaluationen sind, desto mehr methodologische Voraussetzungen werden verletzt; je mehr sich Evaluationen an ein methodologisch sauberes Experimentaldesign unter Laborbedingungen annähern, umso „gültiger“ sind die Ergebnisse auf der Theorieebene, zugleich aber umso unbrauchbarer“ (Bauer:2007:41).

3.3 Vorher – Nachher – Messungen

- Wie war eine Situation, ein Verhalten, eine Einschätzung vor einer Maßnahme und wie ist sie nach der Maßnahme?
- ...verdeutlichen Effekte, die vermutlich durch die Maßnahme hervorgerufen worden sind
- vergleichsweise einfacher zu handhaben, können aber auch komplexer angelegt werden
- dann werden nicht nur Effekte sondern auch noch besonders förderliche Wirkungsfaktoren evaluiert
- werden in nur einer Gruppe vorgenommen, Vergleichs-, Kontrollgruppen entfallen
- Einschätzungen der Wirkungen durch Programmteilnehmer
- Selbsteinschätzung der Adressaten
- Selbsteinschätzung bezüglich Wirkung und Zufriedenheit ist nicht zu unterscheiden
- Zufriedenheit ist kein Kriterium für Wirkung
- ist eher eine Meinungsumfrage und unterhalb des Evaluationsverfahrens anzusiedeln

„Vorher-Nachher-Messungen bei einer Personengruppe, mit der ein bestimmtes Handlungsprogramm realisiert wurde, sind unter den pragmatischen Gesichtspunkten ihrer Realisierbarkeit im Alltag und eines angemessenen Aufwands wohl das am ehesten umsetzbare Design zur Evaluation von Wirkungen in der Sozialen Arbeit“ (Merchel:2010:129).

- Man sieht eine Ausgangssituation, die auf eine begründbare Weise „defizitär“ erscheint
- Dies löst einen Handlungsbedarf aus, auf den reagiert wird
- Ausgangssituation (Vorher) -> Intervention -> veränderte Situation (Nachher)

3.4 Evidenzbasierte Programme

In der Sozialen Arbeit werden zunehmend Arbeitsformen und Programme entwickelt, die systematisch überprüft werden. Wenn diese Programme ihre Wirksamkeit durch entsprechende Studien nachgewiesen haben, müssen bei der späteren Umsetzung keine wirkungsorientierten Evaluationen mehr durchgeführt werden, sondern es muss durch die Evaluation nachgewiesen werden, dass das Programm entsprechend den Standards umgesetzt wurde.

Hilfreich ist die Differenzierung von Evaluationsstandards nach der „Grünen Liste Prävention“, die folgende Stufungen vorsieht:

- Effektivität theoretisch gut begründet
- Effektivität wahrscheinlich
- Effektivität nachgewiesen

(vgl. <http://www.gruene-liste-praevention.de/nano.cms/datenbank/information>)

4 Literatur

Bauer, K.-O. (Hrsg.) (2007): Evaluation an Schulen. Theoretischer Rahmen und Beispiele guter Evaluationspraxis Weinheim/München

Merchel, Joachim (2010): Qualitätsmanagement in der Sozialen Arbeit.: Eine Einführung. Weinheim/München

ISA Planung und Entwicklung GmbH (Hg.)(2009): Band 9: Praxishilfe zur wirkungsorientierten Qualifizierung der Hilfen zur Erziehung, Münster