

IMPACT

PRAXIS

KONZEPT FÜR EVALUATIONEN VON PROGRAMMEN UND ANGEBOTEN

OUTCOME

OUTPUT

WISSENSCHAFT



Konzept für Evaluationen von Programmen und Angeboten

Marie Meierhofer Institut für das Kind

Patricia Lannen, August 2018





Inhaltsverzeichnis

1	Einleitung	5
2	Ziel und Zweck des Konzepts	5
3	Bausteine des Konzepts	6
3.1	Partizipativ - mit und für die Praxis	6
3.2	Das Kind im Zentrum	10
3.3	Multi- und Transdisziplinarität	10
3.4	Multimethodaler Ansatz	11
4	Wirkungsmodelle	12
4.1	Ein erweitertes Modell	12
4.2	Erarbeitung des Modelles	15
4.3	Stolpersteine	17
5	Wahl des Evaluationsdesigns	20
6	Wissensmanagement	26
7	Anhang	28



1. Einleitung

Der Evaluationsansatz des MMI setzt eine systematische Vorgehensweise um und widerspiegelt Kernelemente des MMI Leitbildes (Anhang) sowie des MMI Forschungskonzeptes (Anhang). Es setzt diese in Verbindung mit einer systematischen Analyse und Dokumentation des Feldes und stellt eine Synthese mit Erfahrungen und Expertise aus vergangenen Evaluationen dar.

Das Konzept ist aus mehreren Bausteinen zusammengesetzt und illustriert Werkzeuge, welche im Rahmen der MMI- Evaluationen eingesetzt werden.

Im letzten Kapitel wird illustriert, wie die Nutzung der Erkenntnisse aus den Evaluationen im Sinne von Wissensmanagement optimiert wird.

Der Ansatz stellt hohe Anforderungen an die Forschenden sowohl in Bezug auf deren Fachwissen und die Verfügbarkeit eines breiten Werkzeugkoffers an sozialwissenschaftlichen Methoden als auch auf ihre Erfahrung mit Programmentwicklung und -umsetzung.

2. Ziel und Zweck des Konzepts

Das Evaluationskonzept verfolgt Ziele auf verschiedenen Ebenen:

MMI intern

Das Evaluationskonzept macht das kollektive Wissen und gemeinsame Verständnis zur Umsetzung von Evaluationen explizit, vereinheitlicht so den Wissensstand innerhalb der Organisation und macht diese Basis für alle (neuen) Fachpersonen verfügbar. Es ist ein Qualitätssicherungsinstrument und kann im Sinne von Wissenstransfer und -management kontinuierlich weiterentwickelt werden.

Praxis

Das Konzept schärft die Wahrnehmung für eine zu besetzende Nische im Feld und stärkt die Stellung des MMIs als Kompetenzzentrum für Evaluationen im Transmissionsriemen Forschung, Praxis, Gesell-

schaft. Es bildet eine Grundlage dafür, sich proaktiv in den Diskurs zu Wirkungsorientierung und Wirkungsmessung von Angeboten einzubringen. Das Evaluationskonzept ist ein Werkzeug für Capacity Strengthening hin zu evidenz-basierter Praxis und das Fundament für entsprechende Aufträge, Projekte, Workshops, Schulungen und Publikationen.

Wissenschaft

Das Evaluationskonzept bildet den Ausgangspunkt, sich aktiv an der Stärkung der Evidenzbasis und des Wissens um wirksame Interventionselemente zu beteiligen. Zudem soll es den Austausch über sozialwissenschaftliche Disziplinen hinweg anregen und die Relevanz praxisbezogener Forschung positionieren.

3. Bausteine des Konzepts

Die nachfolgenden „Bausteine“ widerspiegeln die Kernelemente des MMI Leitbildes wie auch des MMI Forschungskonzeptes und beschreiben deren Anwendung auf die Evaluationstätigkeit des MMI. Sie bilden den Kern des Konzeptes.

3.1 Partizipativ - mit und für die Praxis

In diesem Baustein spiegelt sich die partizipative Grundhaltung des MMI. Er veranschaulicht zudem die Arbeitsweise des MMI im Transmissionsriemen zwischen Forschung und Praxis.

Ein partizipatives Vorgehen und eine starke, funktionierende Partnerschaft zwischen dem Praxisanbieter („implementer“) und den Forschenden am MMI, welche von Vertrauen getragen ist, stehen im Zentrum des Evaluationsansatzes. Dafür werden die Ziele der Evalua-

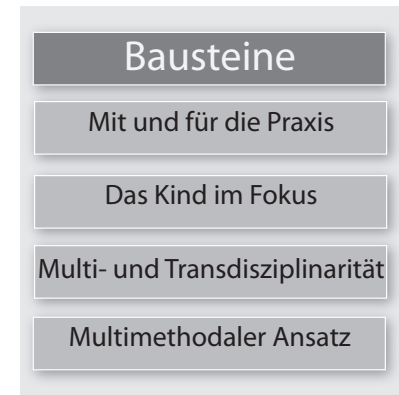


Abbildung 1:
Bausteine des Evaluationskonzeptes

tion, die Kernfragestellungen und das Evaluationsdesign bereits gemeinsam in einem Workshop im Rahmen der Erstellung einer Offerte erarbeitet. Damit wird eine optimale Passung zwischen dem Anliegen der Praxis (Umsetzer und Auftraggeber) und der Evaluation angestrebt, um die Durchführung der Evaluation zu sichern und deren Aussagekraft zu maximieren, Missverständnissen vorzubeugen und die Möglichkeiten und Grenzen verschiedener Evaluationsdesigns transparent zu machen (siehe auch Kapitel 4 zu Wirkungsmodellen und Kapitel 5 zu Wahl der Evaluationsdesigns). Angesprochen wird auch, wie die Unabhängigkeit der Forschung gewährleistet wird, und, was allfällige unerwartete oder gar unerwünschte Ergebnisse für den Praxispartner bedeuten und wie er damit umzugehen gedenkt.

In diesem ersten Workshop werden zudem mögliche Bedenken zur Evaluation besprochen (siehe Box 1). Die MMI-Wissenschaftlerin übernimmt dabei die Rolle eines „Research Brokers“. Sie/er ist jemand, der zwischen der wissenschaftlichen und der praktischen Sprache und Herangehensweise übersetzen kann und damit den Transmissionsriemen zwischen Forschung und Praxis in Schwung bringt.

Weiter wird in diesem ersten Workshop das Wirkungsverständnis des Praxispartners bezüglich dem fraglichen Angebot erfragt, damit die passenden Werkzeuge gewählt und später eingesetzt werden können (siehe Kapitel 4 zu Wirkungsmodellen).

Die Resultate der Evaluation werden gemeinsam mit dem Praxispartner vor Fertigstellung des Schlussberichtes kontextualisiert und wo, angezeigt, werden entsprechende Empfehlungen diskutiert. Bei grösseren Evaluationsvorhaben (über mehrere Jahre) werden Zwischenworkshops zur Vorstellung und Diskussion vorläufiger Resultate eingeschoben.

Wissenschaftliche Unabhängigkeit und Integrität sind in jeder noch so partnerschaftlich aufgelegten Evaluation von grösster Wichtigkeit. Sie werden geschützt und, falls nötig, erklärt. Es werden entsprechende Massnahmen getroffen. Interaktionen und informeller Austausch während der Forschungsphase werden unterlassen und strikt unterbunden. Ein offener Ausgang mit verschiedenen möglichen und ggf. unerwarteten Resultaten muss zu jedem Zeitpunkt gewährleistet sein.

Häufige Einwände und Bedenken aus der Praxis und Möglichkeiten, darauf einzugehen:

„Ich weiss, dass mein Angebot wirkt“

Dieser Eindruck mag entstehen, a) da die Teilnehmenden positiv auf das Angebot reagieren, b) aufgrund spezifischer Beispiele von Teilnehmenden, bei denen das Angebot geholfen hat, c) weil die Intuition und professionelle Erfahrung besagen, dass das Angebot gut ist. Forschungsarbeiten konnten jedoch zeigen, dass die intuitive Beurteilung bezüglich Wirksamkeit oft trügt. Es mag sein, dass vor allem Rückmeldungen aus den Angeboten verzerrt sind und vorwiegend von Personen kommen, welche das Angebot mögen. Es ist auch möglich, dass es eine „Vermischung“ gibt und es Ausdruck davon ist, dass die Personen es schätzen, am Angebot teilzunehmen, und nicht unbedingt, dass es wirkt. Ein Beispiel hierfür ist DARE, ein Drogenpräventionsprogramm in den USA, welches über viele Jahre hinweg umgesetzt, aber erst nach vielen Jahren evaluiert wurde. Es zeigte sich, dass das Programm nicht wirksam ist, und Abschreckung nicht als Wirkfaktor für eine erfolgreiche Prävention ausreicht. Es zeigte sich, dass vielmehr Peer-Einflüsse ausschlaggebend sind und in einer Intervention berücksichtigt werden müssen.

Eine Evaluation kann die intuitive Einschätzung aber auch bestätigen und untermauern, was ein starkes Argument für die Weiterführung und ggf. die entsprechende Finanzierung ist.

„Evaluation ist teuer und nimmt viel Zeit in Anspruch. Sollten die Ressourcen nicht lieber den Kindern direkt zugute kommen?“

Dieses Argument ist sehr gut nachvollziehbar: es ist schwer, Geld zu finden für Evaluationen, und es ist noch schwerer, wenn Gelder zur Umsetzung des Angebotes für diese Kinder und Familien selber auch limitiert sind. Gleichzeitig ist es problematisch und unter Umständen unethisch, eine Dienstleistung anzubieten, von der man nicht weiss, ob sie den gewünschten Effekt überhaupt hat oder im Gegenteil vielleicht sogar Schaden anrichtet.

Wieder ist das Beispiel DARE illustrativ: es wurden über eine Milliarde USD in das Programm investiert, bevor zu Tage kam, dass das Programm keine Wirkung zeigt, d.h. die Adressat/innen ohne passende Unterstützung blieben.

„Was wir tun, ist nicht linear und nicht messbar“

Es ist in der Tat so, dass das Messen von menschlichen und zwischenmenschlichen Phänomenen durch die Sozialwissenschaft oft nicht so offensichtlich und geradlinig ist wie teilweise in anderen Disziplinen. So kann man zum Beispiel in der Medizin häufig mit relativ einfachen Tests das Vorhandensein einer Krankheit definitiv im Blut nachweisen (ja, die Krankheit ist vorhanden oder nein, sie ist nicht vorhanden). Oft ist dies in der Medizin und erst recht in den Sozialwissenschaften nicht ganz so eindeutig und schwarz-weiss. Ein Beispiel ist die Erfassung von Gewalt: Welche Handlungen gelten als Gewalt und welche nicht? In welche Dynamik sind sie eingebettet? Verzerrt die soziale Erwünschtheit eventuell Antworten?

Die Sozialwissenschaften haben in den letzten Jahren enorme Fortschritte hin zu valider und reliabler Messung von Phänomenen gemacht. Ein Evaluationsansatz, der qualitative und quantitative Methoden verbindet, ermöglicht die gleichzeitige Erfassung des "big picture" wie auch der Reichhaltigkeit der Erfahrungen. Sie ermöglicht einen Überblick und spezifische Einblicke in die Tiefe.

„Was mache ich, wenn die Evaluation zeigt, dass mein Angebot nicht funktioniert?“

Evaluationsergebnisse sind in den seltensten Fällen einseitig positiv oder negativ. Oft zeigt sich, dass eine Intervention für gewisse angestrebte Outcomes wirksam ist und für andere weniger. Oft liefert die Evaluation wertvolle Hinweise zur Stärkung und Verbesserung der Intervention. Und wenn die Intervention nicht wirkt oder teilweise am Ziel vorbei geht. Würde man das nicht wissen wollen?

„Es ist unethisch, wenn nicht alle vom Programm profitieren können“

Diese Bedenken werden manchmal für Evaluationen mit Kontrollgruppe (quasi-experimentelle Studien oder randomisierte Kontrollstudien) geäußert und sind in gewissen Fällen gerechtfertigt.

Es ist allerdings zu bedenken, dass oft so oder so nicht genügend Ressourcen vorhanden sind, um alle potentiellen Teilnehmer im Programm einzuschliessen. Die Erhebung von Informationen bei Personen, die nicht am Programm teilnehmen können, ist dann eine Möglichkeit, die letztlich auch der Programmausweitung zugute kommen kann.

Wenn eine Staffelung der Aufnahme in ein Programm nötig und möglich ist, stellt die Einführung von Warte-Kontrollgruppen eine sehr sinnvolle Option dar. Zudem ist zu bedenken, dass das Anbieten einer Dienstleistung, von der wir nicht wissen, ob sie wirkt oder gar negative Effekte hat, ethisch mindestens ebenso bedenklich wie der beschränkte Zugang zum Programm.

„Ich habe die Expertise nicht, eine Evaluation durchzuführen“

Dies ist häufig wahr. Die Ausbildung und die Fertigkeiten eines guten Praktikers/einer guten Praktikerin sind nicht dieselben wie die eines guten Evaluators/einer guten Evaluatorsin. Und je komplexer und im wissenschaftlichen Sinne strenger die Evaluation ("rigour") ist, desto spezialisierter ist das notwendige Wissen. Eine starke Partnerschaft zwischen ForscherIn und PraktikerIn ist deshalb das A und O, also der entscheidende Gelingensfaktor einer erfolgreichen Evaluation.

Quelle: Lannen, P., Jones, L., (in prep). Beyond Evaluation: Using Science to Accompany Program Development and Implementation. Chapter 2: Concerns about engaging in evaluation.

3.2 Das Kind im Zentrum

Das Kind und seine Lebenswelten stehen in der Arbeit des Marie Meierhofer Institutes im Zentrum jeder Tätigkeit. Thematisch werden Evaluationen priorisiert, welche auf die frühe Kindheit oder kinderrechtliche Fragen (vor allem auch im Sinne von Unterstützung von Kindern aus vulnerablen Umständen) fokussieren.

Die Grundhaltung des MMI, „vom Kind her“ und seiner Erlebniswelt zu denken, ist für die Evaluationen des MMI zentral (vgl. Modell des MMI zur Partizipation junger Kinder: Marie Meierhofer Institut für das Kind 2016, Modell kindlicher Partizipation aus Kinder- und aus Erwachsenenperspektive, Poster, Beilage und Kinder 98). Um die Perspektive der Kinder direkt abzuholen, werden Gefässe wie Kinderworkshops und passende Methoden der Kommunikation mit Kindern eingesetzt. Wo immer möglich und zielführend, werden Outcomes auf Ebene Kinder als zentrale Indikatoren eingeschlossen.

3.3 Multi-und Transdisziplinarität

Das MMI beherbergt unter einem Dach Forschungs- und Praxisexpertise (Transdisziplinarität) und zeichnet sich durch eine multidisziplinäre Arbeitsweise mit Fachpersonen aus der Entwicklungspsychologie, Frühpädagogik und Familiensoziologie aus. Zudem arbeitet das Team regelmässig mit Kolleg/innen der Entwicklungs pädiatrie und des Familien- und Kinderrechts zusammen. Dies ermöglicht einen ganzheitlichen Blick auf das Kind sowie das niederschwellige Hinzuziehen eines breiten Spektrums an Expertise, einschliesslich erweitertem Grundlagen- und Praxiswissen in den verschiedenen Disziplinen, Expertise in therapeutischen und pädagogischen Interventionen sowie Organisationsentwicklung. Dies wird für jede Evaluation nach Bedarf entweder informell oder im Rahmen von bereichsübergreifenden Werkstätten hinzugezogen.

3.4 Multimethodaler Ansatz

Der Evaluationsansatz des MMI zeichnet sich durch ein systematisches Vorgehen und durch den Einsatz von Methoden gemäss höchsten sozialwissenschaftlichen Standards aus. Die Mitglieder des Forschungsteams sind in einer der relevanten Methodenrichtungen vertieft geschult (siehe Abschnitt 3.3) und verfügen über eine breit angelegte Tool-box an sozialwissenschaftlichen Methoden. Dies ermöglicht es, je nach Fragestellung das passende Design einzusetzen (siehe Kapitel 5 Wahl des Evaluationsdesigns).

Das Team arbeitet explizit mit qualitativen wie auch quantitativen Forschungsmethoden, wenn immer möglich, ergänzend in jeder Evaluation. Daten werden mit standardisierten und wo, immer möglich, validierten Erhebungsinstrumenten erhoben. Bei massgeschneiderten und innovativen Instrumenten wird hoher Wert auf systematische, reflektierte Anwendung gelegt. Es wird, wie üblich, unterschieden zwischen hypothesen-testendem und explorativem Vorgehen und es werden mehrere Perspektiven zu einer Fragestellung eingeholt (mult-informant). Die Wahl eines Evaluationsdesigns soll explizit nicht nach den (qualitativen vs. quantitativen) Vorlieben des Wissenschaftlers/der Wissenschaftlerin gesteuert, sondern von der Programmreife und den Fragestellungen abhängig gemacht werden (siehe Kapitel 5 Wahl des Evaluationsdesigns).



4 Wirkungsmodelle

4.1 Ein erweitertes Modell

Ein zentrales Werkzeug im Rahmen der MMI Evaluationen ist die gemeinsame Erarbeitung eines Wirkungsmodells ("program theory", "theory of change", "logic model"). Ein Wirkungsmodell verdeutlicht, was die Programmverantwortlichen im Rahmen des Angebotes als notwendig und sinnvoll erachten, und erläutert in logischen Schritten, wie das Angebot das soziale Problem angeht, wie es zu Veränderung führt und was es genau erreichen will. Es benützt bestehende Evidenz, um die einzelnen Schritte zu untermauern, und benennt die zu erreichenden Ziele als klar definierte Outcomes. Es ermöglicht die Identifikation von «blinden Flecken» und Verbesserungspotenzialen und stärkt so bereits das Angebot in Richtung Wirkungsorientierung. Zudem ist ein Wirkungsmodell ein äusserst nützliches Werkzeug, um prägnant zu kommunizieren, worum es im Angebot geht. Es ist die Grundlage jeder MMI-Evaluation.

In vielen Angeboten ist das Wirkungsverständnis entweder intuitiv, nicht explizit zugänglich, hat logische Lücken oder es ist nicht mit der entsprechenden Evidenz untermauert.

Des Weiteren sind Outputs und Outcomes oft vage formuliert, sodass eine Wahl der zu messenden Ziele erschwert ist oder zu Missverständnissen führen kann.

Viele der klassischen Ansätze der Erläuterung von Wirkungsmodellen beginnen direkt bei den Elementen des Angebotes und machen explizit, wie diese zu spezifischen Zielen (outputs und outcomes führen).

Hier zwei Beispiele:

Beispiel 1:

Quelle: Stiftung ZEWO

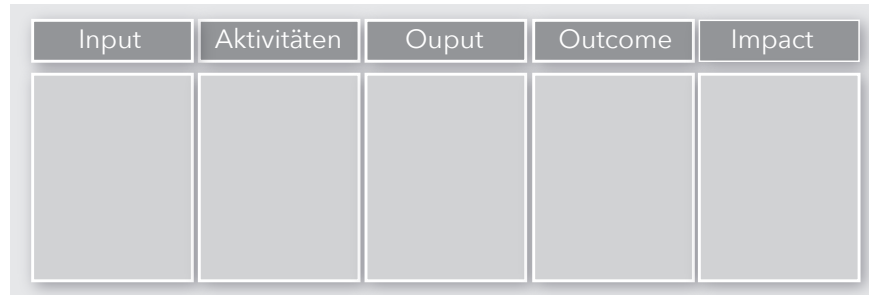


Abbildung 2: Beispiel Wirkungsmodell 1

Beispiel 2:

Quelle: Jones, L. (2014). Improving efforts to prevent children's exposure to violence- a handbook to support child maltreatment prevention programs. World Health Organization.



Abbildung 3: Beispiel Wirkungsmodell 2

Diese berücksichtigen aber meist nicht, welches „soziale Problem“ das Angebot zu lösen versucht („weshalb es das Angebot braucht“) und welche Evidenz zum Problem vorhanden ist (Prävalenzen, Ursachen, Risikofaktoren etc). Dabei sind eine logische Verkettung sowie eine Verlinkung zwischen der Problemstellung und den Angebotselementen sowie wiederum eine Einbettung der Verlinkung mit bestehender Evidenz essentiell.

Das MMI-Konzept erweitert daher das Modell und fügt diese Elemente hinzu: Was soll konkret mit dem Angebot angegangen werden und wie steht es mit der Passung der Programmelemente in Bezug auf das Adressieren des sozialen Problems? Welche Evidenz haben wir zu den Gründen, die zum sozialen Problem führen? Sind sie ausreichend im Angebot berücksichtigt? Was sind die Hindernisse, dass das Problem nicht schon lange gelöst ist (z.B. Zugang zu Zielgruppe), und wie wird dies im Angebot überwunden?

So ergibt sich das folgende Modell: Abbildung 4

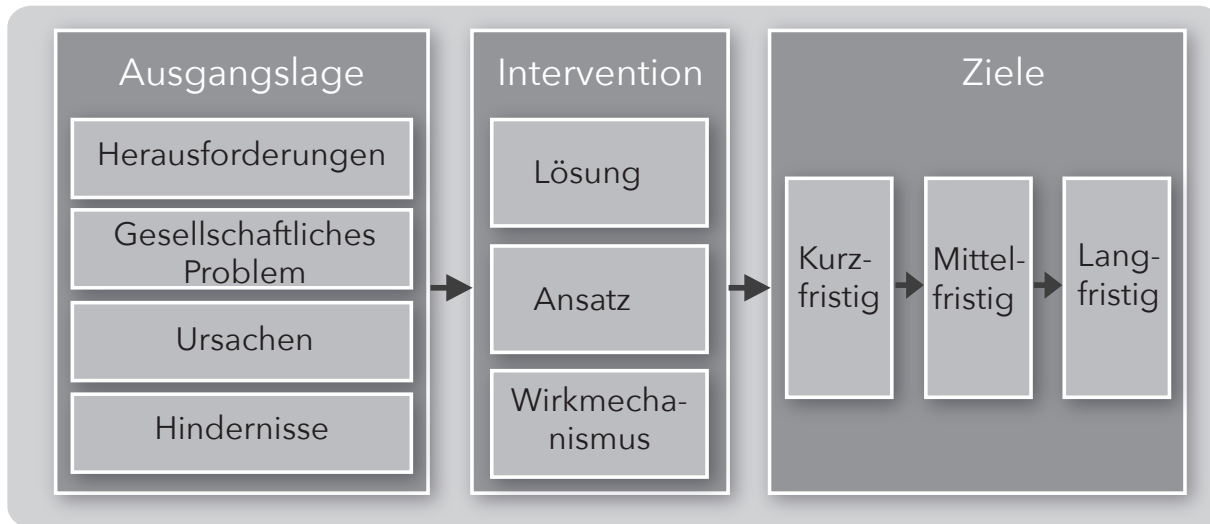


Abbildung 4: MMI- Wirkungsmodell

Erläuterungen:

Ausgangslage: Sie beschreibt das gesellschaftliche Problem, welches das Angebot anzugehen versucht, sowie dessen Ursachen. Sie dokumentiert Hindernisse (warum wurde das Problem nicht schon lange gelöst?) oder spezielle Herausforderungen bezüglich Umsetzbarkeit eines Angebotes (zum Beispiel Erreichbarkeit der Zielgruppe).

Intervention: Die Beschreibung der Intervention soll eine Lösung des identifizierten Problems sein. Der Ansatz soll die Hindernisse und Herausforderungen lösen (z.B. Niederschwelligkeit, um Erreichbarkeit zu verbessern) und den Wirkmechanismus, wie das Problem gelöst wird, explizit machen.

Ziele: Ziele sollen als kurz-, mittel- und langfristige Ziele formuliert und direkt als Lösung für das initial definierte „Problem“ erkennbar sein.

4.2 Erarbeitung des Modelles

Im Rahmen des ersten Workshops wird das Wirkungsverständnis des Praxispartners erfragt (siehe 3.1.). Je nach dessen Differenziertheit und Reifegrad, wird im Rahmen der Evaluation mehr oder weniger Zeit auf dieses Element verwendet. Bei nur rudimentär ausgereiften Wirkungsmodellen wird in einem bis zwei weiteren Workshop gemeinsam und im Dialog ein Wirkungsmodell erarbeitet.

Dazu werden unter anderem die folgenden Leitfragen eingesetzt, um die einzelnen Elemente des Wirkungsmodelles einzufüllen und in logischen Schritten schematisch abzufassen:

- Welche Inhalte werden angeboten?
- Weshalb braucht es das Angebot /was ist das Problem?
- Warum wurde dieses Problem nicht schon lange gelöst (Hindernisse)?
- Wie geht das Angebot das Problem an?
- Wer ist die Zielgruppe des Angebotes und warum?
- Welche Ziele gibt es kurz-, mittel- und langfristig?

- Welche Wirkungen hat das Angebot bei der Zielgruppe?
- Wie werden mit dem Angebot die Ziele erreicht?

Dabei ist die Notwendigkeit des Individualisierungsgrades der Unterstützung meist hoch. Oft sind Praktiker sehr in der operativen Arbeit vertieft und es fällt ihnen schwer, ohne Strukturierungshilfe von aussen das Angebot schematisch abzufassen. Ein selbstständiges Erarbeiten (zum Beispiel mit Leitfragen ohne Dialog) hat sich nicht als erfolgreich herausgestellt.

Wichtig anzumerken ist, dass die im Wirkungsmodell formulierten Ziele nicht unbedingt vollständig deckungsgleich sein müssen mit dem, was dann in der Evaluation erhoben und erfasst wird. Je nach Rahmenbedingungen und Ressourcen der Evaluation, ist es meist nicht möglich, alle Ziele zu messen. Das Wirkungsmodell hilft aber, explizit zu machen, was innerhalb der erwarteten Wirkung und der angestrebten Angebotsziele im Rahmen einer konkreten Evaluation tatsächlich gemessen wird und was nicht. In der Evaluation nicht gemessene Ziele sollen, wenn immer möglich, durch bestehende Forschungsergebnisse aus anderen Studien belegt werden.

Neben der Wichtigkeit des Wirkungsmodells für das Angebot und als Basis für die Evaluation selber, dient die gemeinsame Erarbeitung des Modells als Instrument, um eine Arbeitsbeziehung aufzubauen, die Prioritäten der anderen Partei zu verstehen und Vertrauen zu schaffen, dass wirklich das gemessen wird, was im Zentrum des Angebots und der Fragen der Praktiker steht. Es ist eine Gelegenheit für den Forscher, das Wirkungsverständnis des Praxispartners zu erfassen, zu schärfen und explizit zu machen. Nur so kann die Logik der Impact Kette vom Problem zur Lösung überprüft werden und man kann schauen, ob die Logik der Kette dem, was wir heute aus der Forschung wissen, standhält.

Die Erarbeitung von Wirkungsmodellen setzt die Fähigkeit der Forscherin voraus, hochstrukturiert zu denken, und bedarf einer hohen Abstraktionsfähigkeit, um ein Angebot schematisch darstellen zu können. Zudem braucht es substantielles Fachwissen, um die inhaltlich logischen Schritte nachzuvollziehen und sie mit relevanter Forschungsliteratur zu untermauern. Die Forscherin muss den Praxispartner flexibel abholen und sich auf seine Gedankengänge und Erfahrung einlassen können (niederschwellig und hoch professionell).

4.2 Stolpersteine

Im Folgenden werden drei der häufigsten Stolpersteine bei der Erarbeitung eines Wirkungsmodells mit einem möglichem Umgang damit im Rahmen von MMI-Evaluationen aufgezeigt.

In Zusammenarbeit der Autorin mit der University of New Hampshire und der UBS Optimus Foundation wurden drei kurze Videos entwickelt, welche sich sehr gut als Werkzeuge eignen, mit Praktikern in Dialog zu kommen und einige dieser Punkte wieder aufzunehmen.

Why Evaluation: http://www.unh.edu/ccrc/evaluation-videos/Chapter_1_Master.mp4

Program Theory: http://www.unh.edu/ccrc/evaluation-videos/Chapter_2_Master.mp4

Measuring Outcomes: http://www.unh.edu/ccrc/evaluation-videos/Chapter_3_Master.mp4

Outputs vs. Outcomes

Eine der grossen Herausforderungen ist das Festlegen von zu erreichenden Outcomes und deren Abgrenzung von Outputs. Oft wird die Wirksamkeit oder der Erfolg eines Projektes in Outputs ausgedrückt, die automatisch dem Outcome gleichgesetzt werden. Das Erreichen von 50 Kindern und deren Teilnahme am Angebot heisst zum Beispiel noch nicht unbedingt, dass sich ihre Entwicklung als Folge des Angebots auch wirklich verbessern wird.

Outputs: sind Projektaktivitäten und -leistungen wie Anzahl Durchführungen, Kinder, Sitzungen/Workshops etc.

Es ist essentiell zu beschreiben, wie viele Personen erreicht werden, was deren Merkmale sind und welche Leistungen im Rahmen des Angebotes erbracht werden. Bereits dies stellt Organisationen oft vor grosse Herausforderungen. Im Rahmen einer MMI-Evaluation wird, wenn nötig, immer auch besprochen, wie die wichtigsten Outputs im Rahmen eines Monitoringkonzepts definiert, sinnvoll festgehalten und verwendet werden können.

Outcomes: sind die Veränderungen, welche in der Zielgruppe angestrebt werden.

Schwierigkeiten bereitet in sehr vielen Fällen die sinnvolle Definition von Outcomes. Oft sind sie vage

und nicht oder nur sehr schwer überprüfbar (z.B. „Kinder sind sicherer in der Gesellschaft“; besser wäre es, was gemeint ist, präzise auszuformulieren: „Kinder erleben weniger Gewalt“; „Kinder haben weniger Unfälle“).

Ein häufiges Beispiel einer zu kurzichtigen, nicht zu Ende gedachten Wirkung und vernachlässigter Evidenz ist es von verbessertem Wissen auf eine angestrebte Wirkung auf der Verhaltensebene zu schliessen. Der Transfer von Wissen zu Verhaltensänderungen ist jedoch oft nicht ohne weitere Unterstützung gegeben.

Die klare Definition und ein gemeinsames Verständnis von Outcomes sind eine der wichtigsten Grundlagen für eine erfolgreiche Evaluation. Nur so kann sichergestellt werden, dass im Rahmen der Evaluation auch wirklich das gemessen wird, was das Angebot erreichen soll, und nicht etwas anderes. Ansonsten ist es nicht möglich, fehlende Wirkung (Null-Effekte) sinnvoll zu erklären. In diesem Fall misst der Forscher etwas, das gar nicht vom Angebot intendiert wird. Wenn durch die Messung dann keine Veränderung angezeigt wird, das Angebot also bezüglich dieses Outcomes „nicht wirkt“, ist dies ein Resultat, das keine schlüssige Interpretation ermöglicht.

Leider ist es so, dass dies immer wieder vorkommt und vor allem auf Seiten der Praktiker zu grosser Frustration, Verunsicherung und möglichem Reputationsverlust führen kann.

Logische Lücken

Logische Lücken im Wirkungsmodell zeigen sich meist auf eine dieser zwei Arten:

1. Es zeigt sich, dass die Intervention nicht die effizienteste/beste Art und Weise ist, das Problem anzugehen, oder dass es Forschungsergebnisse gibt, welche besagen, dass zusätzliche Angebots-elemente notwendig sind.
2. Es zeigt sich, dass es nicht möglich ist, mit den zur Zeit umgesetzten Angebotselementen die angestrebten Ziele zu erreichen.

Manchmal stellt sich sogar bei diesem frühen Evaluationsschritt heraus, dass die Diskrepanz entweder zum skizzierten Problem oder den zu erreichenden Zielen so gross ist, dass es sich lohnt, zuerst das



Angebot zu überarbeiten, um die Passung zwischen Problem und Erreichung der Ziele zu optimieren, bevor weitere Schritte der Evaluation umgesetzt werden.

Organisationsentwicklung

Manchmal stellt sich auch heraus, dass das Hauptanliegen einer Organisation nicht programmatischer Art ist, sondern dass Organisationsentwicklungselemente prioritär anzugehen sind. Dies können Unklarheiten in operativen Prozessen sowie bzgl. Verantwortung und Steuerung (Governance) oder mangelnde Finanzierungs- bzw. Fundraisingexpertise sein. Oft sind kleinere Organisationen von Fachpersonen geführt, welche inhaltlich reichhaltige Expertise mitbringen, aber nur minimal geschult sind in Organisationsführung. Bei erfolgreichen Programmen und somit wachsenden Organisationen ist dieses Manko besonders kritisch.

Damit sich die Investition in eine Evaluation lohnt und deren Ergebnisse nicht durch Mängel der Organisationsführung verzerrt werden, kommt es auch in solchen Fällen vor, dass die Forscherin des MMI vorschlägt, zuerst oder zumindest parallel auf einen Organisationsentwicklungsprozess zu fokussieren, bevor die Evaluation des Angebots zu weit fortschreitet. Zur Unterstützung kann das MMI hier auf die interne Organisationsentwicklungsexpertise des pädagogischen Teams und auf ein Netz von externen Berater/innen zurückgreifen.

5. Wahl des Evaluationsdesigns

Die Wahl des passenden Evaluationsdesigns steht neben der Investition in eine solide Partnerschaft ebenfalls im Zentrum von MMI-Evaluationen. Wie unter 3.4. bereits erläutert, werden Evaluationsdesigns nicht aufgrund von Fertigkeiten und Vorlieben des Evaluators (eher qualitativ vs. quantitativ, eher formativ vs. summativ, eher explorative vs. Wirkungsevaluation) gewählt, sondern in einem systematischen Prozess erarbeitet, der verschiedene Faktoren berücksichtigt.

Ziele sind dabei, die Stärke („rigour“) der Evidenz, die aus der Evaluation gewonnen werden kann,

zu maximieren und gleichzeitig den Nutzen für das Programm zu optimieren. Das Design wird auf die gemeinsam definierten Ziele und die Kernfragestellungen der Evaluation abgestimmt. So bedürfen explorative und formative Zugänge neben einer guten Basisdokumentation der Outputs eher qualitativer Methoden, während die Stärke der Evidenz signifikant verbessert werden kann, wenn für hypothesentestende Fragestellungen sowie für Wirkungsevaluationen, wo immer möglich, ein quantitatives Vorgehen gewählt wird.

Zudem fließen die Ausgereiftheit des Wirkungsverständnisses (siehe Kapitel 4) sowie die Phase, in der das Angebot umgesetzt wird (Konzeptphase, Pilotphase, Umsetzungsphase, Scale up Phase), wesentlich in die Entscheidung mit ein.

Nicht zuletzt ist selbstverständlich der verfügbare Ressourcenrahmen (finanzielle Mittel und Zeithorizont der Evaluation) ausschlaggebend.

Im Idealfall wird in der Entwicklung eines Angebots bereits in der Konzeptphase ein logisches, explizites Wirkungsmodell erarbeitet, welches mit wissenschaftlichen Erkenntnissen untermauert ist. In einer Pilot- und anschließenden Umsetzungsphase werden fortlaufend systematisch formative Daten erhoben, die Praxis reflektiert und die Umsetzung gefestigt. Zudem wird eine Abfolge zunehmend strengerer Wirksamkeitsevaluationen durchgeführt. Erst wenn die Machbarkeit und Wirksamkeit eines Angebotes etabliert sind, ist es sinnvoll und vertretbar, unterstützt mit weiteren formativen Elementen, eine Verbreitung (Scale up) dieses nun evidenz-basierten Angebots anzustreben.

Ein kontinuierlicher und aktualisierter Einbezug von Forschungsergebnissen und eine auf den Umsetzungsstand angepasste Evaluation des Angebots sind in jeder Phase wertvoll und wichtig. Im Rahmen der Evaluationsarbeit begleitet das MMI Angebote in jeder Phase der Angebotsentwicklung. Intern ist das MMI bestrebt, im Austausch zwischen dem Fach- und dem Forschungsteam die eigenen Angebote auf diese Weise zu entwickeln und kontinuierlich zu evaluieren. Dies im Bewusstsein, dass sich damit je nach Fragestellung und Art des Angebots zu gegebener Zeit eine externe Evaluation nicht erübrigt.

Viele Angebote und Programme werden allerdings nicht in einem solch engen Austausch zwischen

Praxis und Forschung entwickelt. Die Realität der meisten Evaluationsaufträge, welche an das MMI gelangen, sieht daher etwas anders aus. Viele Praktiker haben keinen oder limitierten Zugang zu wissenschaftlichem Vorgehen und/oder werden während der Programmentwicklung nicht von einem wissenschaftlichen Partner begleitet. Deshalb gelangen Anfragen für einen externen Evaluationsauftrag in jeder Angebotsphase und mit jeder Art von Ausreifung der Nutzung von Evidenz an uns. Der MMI-Evaluationsansatz schätzt zur Bestimmung des passenden Evaluationsdesigns systematisch einerseits die Projektphase und die „Reife“ des Programms oder Angebots und andererseits die bereits bestehende Evidenzebene ein. Es ergeben sich die folgenden möglichen „Stadien“ von „Angebotszuständen“ (Tabelle 1). Die Bedeutung der Begriffe evidence-informed und evidence-based werden in Box 2 erläutert.

		Projektphase			
		Konzept	Pilot	Umsetzung	Scale Up
Evidence of impact	Keine	1	2	3	4
	Evidence- informed	5	6	7	8
	Evidence- based	9	10	11	12

Tabelle 1: Angebote

Evidence-informed (auf reflektierte Erkenntnisse oder auf Forschungsergebnisse aus anderen Studien gestützt) bedeutet, dass das Projekt in ein Wirkungsmodell eingebettet ist, auf bestehenden Forschungserkenntnissen aufbaut und eine systematische Umsetzung von Praxisreflexion und/oder formativer Evaluation besteht.

Evidence-based (auf belegter Wirksamkeit basierende Evidenz) bedeutet, dass Erkenntnisse zur Wirkung bestehen, welche in einer formalen Wirksamkeitsevaluation durchgeführt wurden, idealerweise in einem Design mit Vergleichsgruppe.

Es ist nicht zu erwarten und auch gar nicht möglich, dass ein Angebot in der Konzeptphase bereits Wirksamkeit nachweisen kann, da noch keine Umsetzung stattgefunden hat (Feld 9). Weiter ist es eher selten und nicht zwingend notwendig, dass ein Angebot in der Pilotphase bereits Wirksamkeit auf einem hohen Niveau (oder überhaupt) nachweisen kann (Feld 10). Spätestens in der Umsetzungsphase sollte aber in die Wirksamkeitsevaluation in Richtung Evidence of impact investiert werden. Es gibt aber Angebote, die bereits weit verbreitet sind und grossflächig umgesetzt werden (Scale up), jedoch in keiner Weise evidenz-basiert sind und deren Wirkung oder potentielle Risiken oder unerwünschte Effekte unbekannt sind (Feld 4). Im Rahmen einer MMI-Evaluation wird zu Beginn der Evaluation besprochen, wie schrittweise die entsprechenden Stufen „aufgeholt“ und die Wirkungsorientierung und Evidenz des Angebotes gestärkt werden können. Um ein Angebot „evidence-informed“ zu machen, werden Wirkungsmodelle erarbeitet und sie in die bestehenden wissenschaftlichen Erkenntnisse eingebettet, formative Evaluationen durchgeführt und ein Konzept für reflektierte Praxis eingeführt. Anschliessend werden in zunehmender Stärke (rigour) der Evaluationen Erkenntnisse zur Wirkung etabliert.

Das MMI setzt explizit keine strengen Evaluationen wie randomisierte Kontrollstudien mit schlecht etablierten und konzeptuell mangelhaft eingebetteten Angeboten um. Ein zu früh eingesetztes, strenges Evaluationsdesign – wenn dies auch für den auf wissenschaftliche Publikationen angewiesenen Forscher interessant ist – schwächt die Chancen auf eine erfolgreiche Durchführung einer Evaluation beträchtlich, schmälert den möglichen Erkenntnisgewinn und birgt damit ein substantielles Risiko für den Praxispartner. Es ist in der Arbeit mit dem Praxispartner sorgfältig zu beachten, dass oft grosse Verwirrung bezüg-

lich Möglichkeiten und Grenzen eines Evaluationsdesigns besteht: ein Angebot, welches sich einer formativen Evaluation unterzogen hat, kann keine Aussagen zur Wirksamkeit des Angebotes machen und es kann sich unter Umständen „evidence –informed“, aber nicht „evidenz-basiert“ nennen. Aus diesem Grund wird im Rahmen des Workshops zur Wahl des Designs mit den Praktikern im Detail diskutiert, was die Vor- und Nachteile eines Designs sind, welche Art von Fragen mit den unterschiedlichen Herangehensweisen beantwortet werden und mit welcher „Sicherheit“ entsprechende Aussagen zu Wirksamkeit gemacht werden können (siehe Box 3).

Prozess- und Wirksamkeitsevaluation:

Unterschiedliche Evaluationsansätze sind zu unterschiedlichen Zeitpunkten und Phasen eines Angebotes nützlich. Jeder Ansatz beantwortet unterschiedliche Fragen und liefert andere Informationen.

Prozessevaluation

Umsetzung

Das wichtigste Ziel bei einer Prozessevaluation ist das Sammeln von Information, welche in die Angebotsumsetzung gespiesen werden kann. Zum Beispiel:

- Wie gut wird das Angebot umgesetzt? Wird so umgesetzt wie es geplant war?
- Wie gut wird die Zielpopulation erreicht?
- Welche Knacknüsse zeigen sich?
- Welche möglichen Lösungen wurden ausprobiert?
- Welche Elemente bewähren sich / werden nachgefragt, welche nicht?

Wirksamkeitsevaluationen

Prä-Post

Hier wird die Ausprägung der Kernoutcomes zu Beginn und zum Schluss einer Intervention gemes-

sen. In manchen Fällen ist ein Prä-Post Design ausreichend, manchmal braucht es aber eine Vergleichsgruppe, um zu bestimmen, ob die gemessene Veränderung zwischen den Zeitpunkten tatsächlich auf das Angebot zurückzuführen ist. Ein Beispiel dafür sind Indikatoren zur Entwicklung des Kindes: Ein Kind reift und entwickelt sich natürlich über die Zeit. Es ist anzunehmen, dass entsprechende Werte zu einem zweiten, späteren Zeitpunkt höher ausfallen als in einer ersten Messung. Dies geschieht sehr wahrscheinlich auch ohne Intervention.

Um zu entscheiden, ob und zu welchem Ausmass dies aber auf die Intervention zurückzuführen ist, braucht es eine Gruppe, welche nicht am Angebot teilnimmt. Möglicherweise zeigen Vergleiche, dass sich der Rückstand zu einer unbelasteten Gruppe verkleinert oder durch das Angebot ein Abstand zu einer Gruppe ohne Intervention entsteht (beides z.B. beim Erwerb der Umgebungssprache von Kindern aus zugewanderten Familien durch den Besuch einer qualitativ guten Spielgruppe).

Mit Vergleichsgruppe

Quasi-experimentell:

Die Strenge der Evaluation und deren Aussagekraft verbessern sich also, wenn eine Vergleichsgruppe zur Verfügung steht. Manchmal kann es "einfach" eine andere Gruppe von Personen sein, in einem regulären Angebot, aus einer anderen Gemeinde. Aber auch dies kann zu Problemen führen, da die Teilnehmenden sich unter Umständen bereits zu Beginn unterscheiden und die Resultate somit verzerrt sind. Deshalb braucht es eine Ausgangsmessung als sogenannte Baseline, also eine Prä- und Postmessung.

Randomisierte Kontrollstudien (RCT):

Dafür gilt es eine Vergleichsgruppe zu finden, die der Gruppe, welche das Angebot erhält, so ähnlich wie möglich ist. Die Auswahl, wer das Angebot erhält und wer nicht, soll per Zufall entschieden werden. Diese Art von Evaluation wird in der Forschung als Goldstandard gehandelt, ist aber nicht ganz einfach durchzuführen. Sie ist sehr ressourcenintensiv.

6. Wissensmanagement

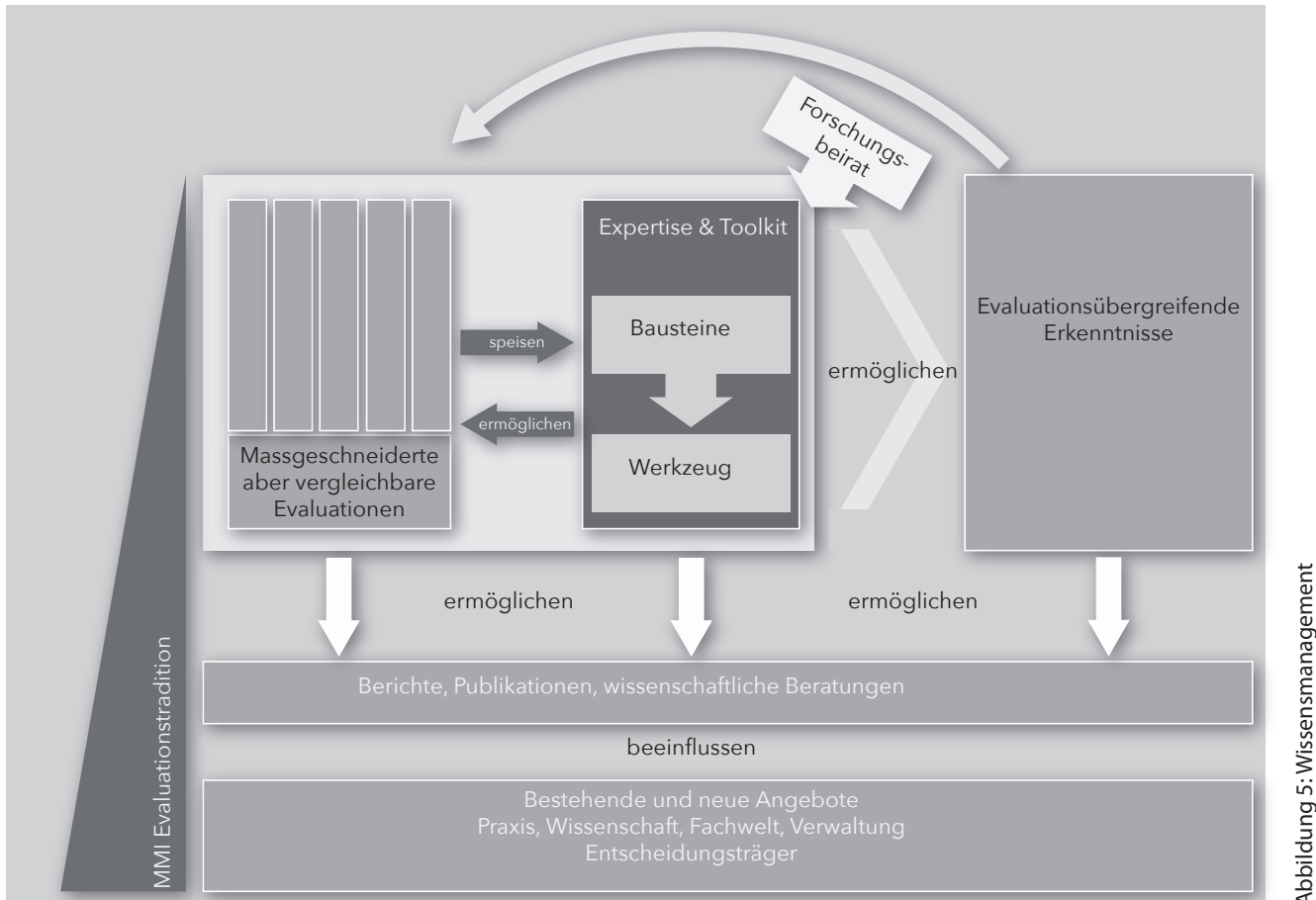
Im Zentrum von MMI-Evaluationen steht schliesslich auch der Anspruch, dass die Resultate der individuellen Evaluationen stets auf mehreren Ebenen einen Erkenntnisgewinn bringen sollen. Um evaluationsübergreifende Erkenntnisse ableiten zu können, ist ein systematisches Vorgehen unabdingbar und bedarf des in diesem Konzept beschriebenen massgeschneiderten und gleichzeitig vergleichbaren und deshalb standardisierten Vorgehens (Abbildung 5).

Im Fokus steht selbstverständlich der Erkenntnisgewinn zu Gunsten des evaluierten Angebots selber. Dieses soll sich anhand der Resultate weiterentwickeln können und zu einem wirkungsorientierten Ansatz gestärkt werden. Dazu werden die Rohresultate gemeinsam mit dem Praxispartner in einem Workshop besprochen und in einem wissenschaftlich soliden, jedoch in erster Linie für Praktiker verständlichen Bericht abgefasst (gray paper).

Die Resultate sollen aber auch anderen, ähnlichen Angeboten zur Verfügung stehen, um kollektives Lernen im Felde zu stärken und die Entwicklung von *good practice* zu unterstützen. Wenn immer möglich, verfasst das Team des MMI Publikationen, in denen diese angebotsübergreifend relevanten Resultate synthetisiert sind. Auf diese Weise sollen die einzelnen Evaluationen wie auch die Evaluationsarbeit in ihrer Summe Relevanz für soziale Veränderungen sowie politische und fachliche Diskussionen gewährleisten.

Wenn immer möglich, sollen die Resultate aus der Evaluation des Angebotes auch in wissenschaftlichen, peer-reviewed Artikeln abgefasst werden, um die Evidenzbasis zu stärken und die Resultate der wissenschaftlichen Gemeinschaft zugänglich zu machen. Ergeben sich Erkenntnisse aus einer Serie von Evaluationen, seien diese inhaltlicher oder methodischer Art, werden diese als Meta-Ergebnisse ebenfalls in wissenschaftlichen Artikeln abgefasst.

Intern werden die Erkenntnisse gebündelt und die Bausteine und Werkzeuge des Evaluationskonzeptes bei Bedarf entsprechend ergänzt oder abgeändert.



7. Anhang

Leitbild Forschung Marie Meierhofer Institut für das Kind

Analog zum Leitbild des Marie Meierhofer Instituts für das Kind beschäftigt sich die Forschung am MMI mit der ganzheitlichen und gesunden Entwicklung des Kindes. Der Schwerpunkt liegt dabei auf der frühen Kindheit. Die Aktivitäten bewegen sich stets am Transmissionsriemen von Forschung, Praxis, Verwaltung und Gesellschaft und verfolgen ein duales Ziel, indem parallel sowohl Forschungsimpact wie auch sozialer Impact angestrebt werden. So wird, um die Relevanz der Erkenntnisse zu gewährleisten, die umgesetzte Forschung massgeblich von Impulsen aus der Praxis, Verwaltung und Gesellschaft gespiesen. Die Resultate werden entsprechend kontextualisiert.

Um das Kind und seine Lebenswelten möglichst ganzheitlich zu erfassen, werden ausgewählte Themen durch ein multi-disziplinäres Fenster beleuchtet. Vertreten sind im Team des MMI namentlich die folgenden Disziplinen: Frühpädagogik (Erziehung und Bildung von Kindern unter sechs Jahren), Entwicklungspsychologie (zeitlich überdauernde und aufeinander aufbauende Veränderungen des kindlichen Erlebens und Verhaltens), Entwicklungspsychopathologie (abweichende Entwicklung und deren Ursachen und Verläufe bei Kleinkindern) und Familien- und Bildungssoziologie (Sozialisation des Kindes in der Familie und den weiteren Lebenswelten). Ferner arbeitet das Team des MMI regelmässig mit Expert/innen der Entwicklungspädiatrie und des Familien- und Kinderrechts zusammen.

In den einzelnen Disziplinen und übergreifend werden Schwerpunkte gesetzt, innerhalb derer dann spezifische Fragestellungen bearbeitet werden.

Ebenfalls wird ein multi-methodaler Ansatz verfolgt. Dies bedeutet, dass sowohl quer-, wie auch längsschnittliche Studien mit qualitativen und / oder quantitativen Methoden umgesetzt werden. Zudem werden auch theoretisch-konzeptionelle Arbeiten durchgeführt.

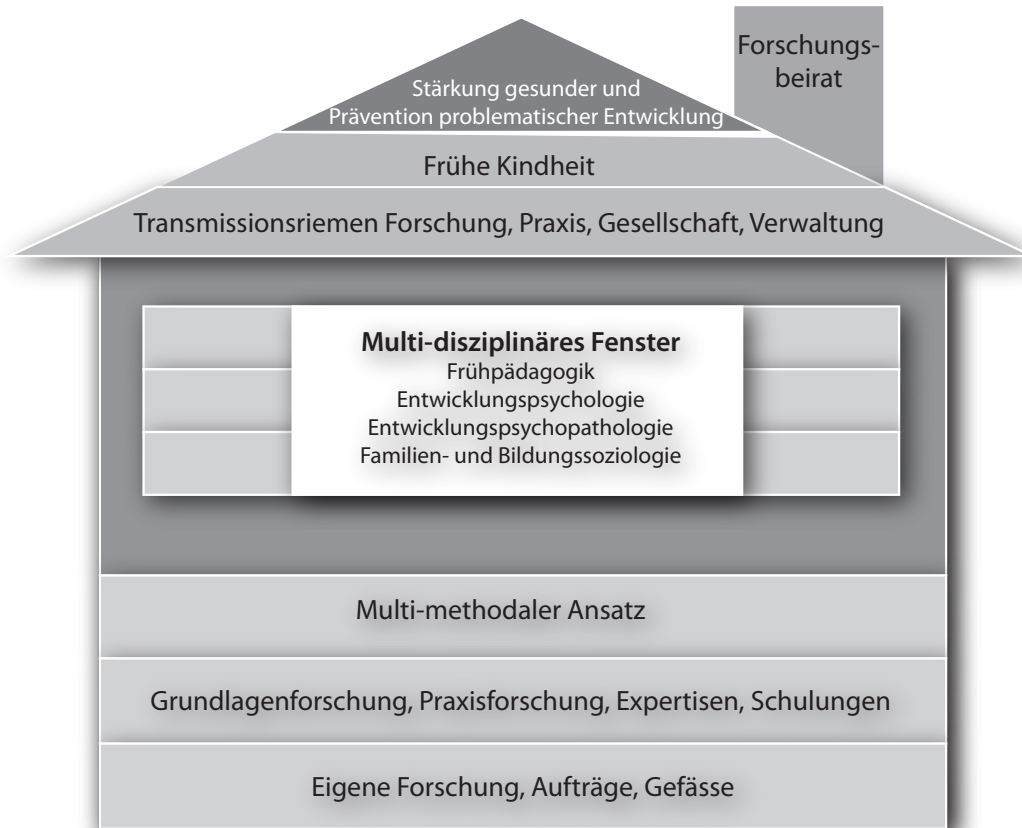


Abbildung 6: Leitbild Forschung

Umgesetzt wird dieser bewusst breit gesetzte Ansatz anhand von strategischer Suche und Anstellung von Personen mit komplementärer methodischer Expertise und unterschiedlichem disziplinären Hintergrund.

Neben Grundlagenforschung, in der erkenntnisorientiertes Elementarwissen erarbeitet wird, hat auch die Praxisforschung einen wichtigen Stellenwert. Diese beinhaltet Evaluationen, Angebots- und Konzeptentwicklung mit wissenschaftlicher Begleitung und die Erarbeitung von Wirkungsmodellen. Es werden sowohl eigene Forschungsideen konzipiert und in Projekten umgesetzt, wie auch Forschungsaufträge Dritter bearbeitet. Des Weiteren bietet das MMI Expertisen und Schulungen als Dienstleistungen an.

Zudem führt das MMI regelmässig Veranstaltungen in einem Spektrum von internen, halb-offenen (mit selektiver Einladung) und offenen (öffentlich ausgeschrieben) Gefässen durch. Es sind dies Werkstätten, Kolloquien und die jährliche wissenschaftliche Tagung zu frühkindlicher Bildungs- und Entwicklungsforschung (in Kooperation mit der Universität Zürich). Die Gefässe dienen dazu, wissenschaftliche Fragen gemeinsam zu diskutieren, gemeinsame Lerngelegenheiten anhand konkreter Projekte zu schaffen und den Transfer von Fragen aus der Praxis in die Forschung und von wissenschaftlichen Erkenntnissen in die Praxis zu gewährleisten.

Die Forschung des MMIs wird durch einen Forschungsbeirat unterstützt. Dessen Funktion ist es, die Forschungsstrategie und die konkreten Forschungstätigkeiten des MMI zu reflektieren, zu diskutieren und beratend zu begleiten.

Herausgeber: Marie Meierhofer Institut für das Kind

Umschlag und Illustrationen: Annalea Guarisco

Layout/Gestaltung: Claudius Natsch

Druck: Borer Druck AG, Laufen

Vertrieb: Marie Meierhofer Institut für das Kind
Pfingstweidstrasse 16, 8005 Zürich
Tel.: 044 205 52 20 / Fax.: 044 205 52 22
info@mmi.ch / www.mmi.ch

© mmi, 2018

