
Susanne Giel

Vom Nutzen der Programmtheorie in Evaluationen im Schulkontext¹

Zusammenfassung

Lehrerinnen, Lehrer und Schulleitungen stehen immer wieder vor neuen Herausforderungen, die gemeistert werden müssen. (Selbst-)Vergewisserung über gelungene, wirkungsvolle Lösungsstrategien ist wichtig; klassische Designs der Bildungsforschung sind jedoch häufig nicht anwendbar, oder ihre Ergebnisse können nicht ohne weiteres in der Schulpraxis genutzt werden. Eine auf Programmtheorien basierende Evaluation versucht, die Lücken zu schließen, wirkungsorientiert und gleichzeitig praxisorientiert vorzugehen. Im Unterschied zu üblichen Schulevaluationen steht dabei nicht die Schule als Organisation, sondern das pädagogische Programm bzw. Handeln auf dem Prüfstand. Der Artikel erläutert die Grundzüge eines theoriebasierten Evaluationskonzepts und illustriert sie am Beispiel einer fiktiven Evaluation von Willkommensklassen. Abschließend reflektiert die Autorin den Nutzen dieses Vorgehens für die Schulpraxis, für das Wissen um Lern- und Bildungsprozesse und auch für methodische Entscheidungen bei Evaluationen.

Schlüsselwörter: Programmevaluation, Programmtheorie, theoriebasierte Evaluation, Innovationen, Wirkungen, Blackbox, Überprüfung von Programmtheorien

The Use of Program Theory in Evaluation in the Context of Schools

Summary

Pedagogic staff in schools consistently faces new challenges. It has to make sure to find succeeding and effective solutions and answers. Classical educational research designs are often either inapplicable, or their results cannot simply be put into practice within the particular context. Theory-based evaluation tries to close the gap by choosing an impact-oriented and simultaneously practice-oriented approach. This article explains the main functions of program theories in evaluations. Unlike typical school evaluation, a theory-based approach puts the educational program, not the school as an organization in its center. The text illustrates the main features of program theories along the notional evaluation of educational efforts to integrate refugee children into the regular school sys-

1 Für die kritischen Kommentare und Rückfragen danke ich herzlich Simone Stroppel.

tem. Concluding, the author reflects the use of this approach for practitioners, for gaining knowledge about learning processes and educational processes, as well as for methodical decisions in the course of evaluations.

Keywords: program evaluation, program theory, theory-based evaluation, innovations, effects, black box, testing program theories

1. Ausgangslage

Bildungsforschung blickt in Deutschland auf eine lange Tradition zurück. Seit dem sogenannten PISA-Schock spielen auch vermehrt Bewertungen eine große Rolle, und somit hat Evaluation an Schulen Einzug gehalten. Im föderalen Bildungssystem wurde dazu eine breite Palette an Verfahren entwickelt (vgl. URL: <http://www.bildungserver.de/Schulentwicklung-Institute-und-Materialien-der-Laender-5079.html>; Zugriffsdatum: 11.04.2016). Diese Verfahren reichen von individuellen Selbstevaluationen, die in erster Linie die Zielerreichung konkreter Unterrichtsvorhaben überprüfen, über externe Peer-Evaluationen bzw. Schulbegehungen, die vor allem Schulentwicklungsprozesse in den Blick nehmen, bis hin zu fächer- und klassenstufenspezifischen Lernstandserhebungen, die Vergleiche zwischen z.B. Schulklassen, Schulen, Bundesländern und Ländern ermöglichen. Insofern kann festgestellt werden, dass es in Schulen eine vielfältige Evaluationspraxis gibt (siehe auch Griese 2016).

Teilweise im Rahmen der Qualitätsdebatte angestoßen, teilweise durch allgemeine gesellschaftliche Entwicklungen und Ereignisse ausgelöst, finden an Schulen innovative Prozesse statt, die Antworten auf aktuelle Herausforderungen liefern sollen. Schlaglichtartig können folgende Handlungsfelder genannt werden: Sprachförderung, interkulturelle Öffnung, Demokratieverziehung und kulturelle Bildung als neue fächerübergreifende Aufgabenstellungen, die Einbeziehung externer Kooperationspartner (insbesondere im Rahmen der Ganztagsbeschulung) und verändertes Lernen durch digitale Medien und das Internet. Aktuell stellt die Integration von geflüchteten Kindern besondere Anforderungen an viele Schulen.

Auf alle diese Herausforderungen müssen Schulen Antworten finden, indem sie Innovationen erproben und sich damit oft auf unbekanntem Terrain bewegen. In solchen Situationen sind Wissen und Erfahrung wertvoll. Evaluation als wissenschaftliche Dienstleistung verfügt über Vorgehensweisen und Methoden, die solche Informationen generieren und die praktisches Handeln wirkungsorientiert unterstützen (vgl. Böttcher/Hense in diesem Heft, S. 117-135). Hierbei sind Antworten auf Fragen danach, was funktioniert bzw. wie neue Handlungsansätze funktionieren, von großem Interesse für die beteiligten Akteure. Die traditionellen Evaluationsstrategien stoßen in solchen Settings jedoch an ihre Grenzen (vgl. dazu ausführlich Giel 2013, S. 55ff.).

2. Grenzen klassischer Forschungsdesigns

Traditionell gelten in der empirischen Bildungsforschung Kontrollgruppenvergleiche als Königsweg, insbesondere dann, wenn es darum geht, die Wirkungen pädagogischen Handelns zu erfassen. Das Grundprinzip besteht darin, mögliche Einflussfaktoren zwischen der Ursache (der pädagogischen Intervention) und der daraus resultierenden Wirkung (bspw. auf Wissen und Kompetenzen) zu kontrollieren, um damit eindeutig den Einfluss der Intervention auf das Resultat zu erfassen. Dies geschieht im Idealfall durch Randomisierung, also eine zufällige Verteilung von Teilnehmenden auf Versuchs- und Kontrollgruppen.

Häufig angewendet werden auch solche Designs, bei denen (Test-)Daten zu verschiedenen Messzeitpunkten – in der Regel vor und nach einer Intervention – herangezogen werden, um Aussagen darüber zu treffen, welche Resultate mit einem neuen Bildungsangebot, einer neuen Lehrmethode etc. erzielt werden können. Beide Vorgehensweisen sind voraussetzungsvoll: Sie bedürfen ausreichend hoher Fallzahlen, bei Kontrollgruppenvergleichen eines Mindestmaßes an Einfluss auf die Gruppenbildung und grundsätzlich der Chance, bereits vor dem Beginn der Intervention zu messen.

Im zuvor geschilderten Szenario, bei dem Neuerungen in den Schulalltag eingeführt werden, ergeben sich weitere Voraussetzungen, die nur selten gegeben sind: Es muss gewährleistet sein, dass die Umsetzung plangemäß stattfindet und dass sie stabil (in vergleichbarer Weise) realisiert wird. Außerdem benötigen sowohl Kontrollgruppenvergleiche als auch bloße Vorher-Nachher-Messungen eine solide Wissensbasis darüber, welche Resultate die Neuerung überhaupt auslösen *kann*, die dementsprechend zu messen wären (vgl. Kromrey 2001).

Auch ein Vorgehen, das sich auf die Überprüfung der Zielerreichung konzentriert, weist deutliche Grenzen auf. Gerade dann, wenn sich Akteure in Schulen mit der Einführung innovativer Prozesse auf Neuland begeben, ist die Formulierung von operationalisierbaren Zielen sehr herausfordernd. Es fehlen Erfahrungen, um abschätzen zu können, was realistisch erreichbar ist und wo die Grenze zwischen Gelingen und Nicht-Gelingen verläuft. Vergleichs- bzw. Kontrollgruppendesigns wie letztlich auch zielorientierte Evaluationen müssen sich darüber hinaus mit dem sogenannten Blackbox-Problem auseinandersetzen (vgl. z.B. Pawson/Tilley 2004, S. 11): Solange ausschließlich der Input und die Resultate betrachtet werden, bleibt der Transformationsprozess dazwischen ausgeblendet. Wodurch nun genau Ergebnisse erzielt bzw. weswegen eben keine Erfolge gemessen werden können, dazu liefern diese Designs keine Hinweise. Lehrerinnen und Lehrer in der Praxis haben dazu zwar häufig Vermutungen; diese werden jedoch nicht explizit und systematisch betrachtet.

Auch die Bedingungen, unter denen Erfolge (nicht) erzielt werden, bleiben von diesen Forschungsdesigns ausgespart.

Zusammenfassend lässt sich sagen, dass die klassischen Herangehensweisen bei der Evaluation von Neuerungen zwei Begrenzungen aufweisen: Zum einen fehlt es noch an Wissen über relevante Dimensionen des Gelingens bzw. des Nicht-Gelingens, auf die man zurückgreifen kann, um dementsprechende Daten zu erheben. Zum anderen bleibt die Brücke zwischen Handeln und Resultaten genauso ausgeblendet wie der Kontext, in dem und für den die Lösungen gesucht werden. Genau hier setzt der im Folgenden vorgeschlagene Evaluationsansatz an, der sogenannte Programmtheorien in seinen Mittelpunkt stellt.

3. Auf Programmtheorien basierende Evaluationen

Theoriebasierte Evaluationskonzepte haben ihren Ursprung in der Evaluation von Programmen, also nicht in der Evaluation von Organisationen (wie bspw. Schulen) oder Produkten (wie bspw. Lehrmedien). Programme werden in der Regel mit Hilfe eines intentional aufeinander bezogenen Sets an verschiedenen Einzelprojekten umgesetzt, die wiederum aus einem Bündel einzelner Maßnahmen bestehen. Ob nun komplexe Programme, fest umrissene Maßnahmen oder einzelne Interventionen – sie werden unter spezifischen Bedingungen geplant und umgesetzt und sind auf spezifische Ziele ausgerichtet (siehe URL: <http://eval-wiki.org/glossar/Programm>; Zugriffsdatum: 11.04.2016). Im Kontext Schule sind aktuelle Programme beispielsweise BiSS („Bildung durch Schrift und Sprache“), das Landesprogramm „Gute gesunde Schule“ in Berlin oder „Kultur und Schule“ in Nordrhein-Westfalen. Programmtheorien – das konzeptionelle Herzstück des im Folgenden dargestellten Evaluationskonzepts – können sich auf komplexe Programme, Projekte und auch einzelne Maßnahmen beziehen.

Im Mittelpunkt der Evaluation stehen das Programm, das pädagogische Handeln, möglicherweise auch einzelne Maßnahmen des Programms (vgl. Balzer/Beywl in diesem Heft, Abschnitt 2, S. 193-197; Böttcher/Hense in diesem Heft, S. 117-135). Explizit nicht evaluiert wird die Organisation Schule. Sie stellt zwar einen Teil der Bedingungen des Programms, wie bspw. die Anzahl und Qualifikationen von Lehrerinnen und Lehrern, vorhandene Räume und deren Ausstattung, die Lehr- und Lernkultur, qualitätssichernde Verfahren etc. Daneben könnten auch der Sozialraum, in dem die Schule liegt, die Eigenschaften von Schülerinnen und Schüler und deren Familien, die eine Schule besuchen, zusätzliche finanzielle Mittel, die zur Umsetzung des Programms zur Verfügung stehen, wichtige Bedingungen des Programms bieten, die sich als hinderlich oder förderlich für die Programmumsetzung und Programm-erfolge erweisen.

Es gilt zu klären, was sich hinter dem Begriff „Theorie“ verbirgt. Ganz allgemein bezeichnet man einen bereits empirisch bestätigten oder noch vermuteten Zusammenhang zwischen mindestens zwei Sachverhalten als Theorie (vgl. z.B. Kromrey 2009, S. 47). In theoriebasierten Evaluationskonzepten wird der Begriff „Theorie“ tatsächlich auch so breit verstanden: Er reicht von Alltagstheorien, die auf persönlichen oder professionellen Erfahrungen beruhen, bis hin zu empirisch überprüften Theorien aus der Forschung oder aus vorangegangenen Evaluationen. Der Begriff Programmtheorie lässt sich in Anlehnung an Bickman (vgl. 1987, S. 5f.) entsprechend definieren als Annahme darüber, in welcher Weise ein Programm Veränderungen herbeiführt. Eine Programmtheorie muss also mindestens geplante Interventionen eines Programms oder einer Maßnahme und die damit auszulösenden Resultate enthalten. Ergänzend sind außerdem – zumindest Pawson/Tilley (2004) betonen das – die relevanten Bedingungen (wie bspw. die Charakteristik der Zielgruppe, eingesetzte Ressourcen etc.) zu explizieren.

Ein auf Programmtheorien basierendes Evaluationskonzept verfolgt das Anliegen, Licht in die Blackbox zu werfen, die traditionelle Evaluationskonzepte oftmals hinterlassen. Die klassischen Fragestellungen von Evaluationen „Wirken Programme?“ oder „Erreichen Programme ihre Ziele?“ erhalten so einen neuen Akzent. Mit diesem Konzept rücken Fragen danach, *wie* Programme wirken, wie sie ihre Ziele erreichen, ins Zentrum der Aufmerksamkeit. Patricia Rogers (2000), eine der wichtigen Protagonistinnen des theoriebasierten Evaluationsansatzes, überschreibt sehr treffend einen ihrer Artikel mit: „Program Theory. Not Whether Programs Work, but How They Work“.

Was bedeutet dies nun für die Evaluation? Zunächst einmal ist festzustellen, dass jede Evaluation über einen theoretischen Kern verfügt. In theoriebasierten Evaluationen wird genau dieser theoretische Kern explizit gemacht und systematisch herausgearbeitet. Dabei werden die Annahmen, die der Evaluation und dem Programm zugrunde liegen, aus dem Programm selbst abgeleitet und nicht – zumindest nicht nur – aus einer wissenschaftlichen Perspektive an das Programm herangetragen. Dadurch wird gewährleistet, dass die Evaluation nicht abstrakt bleibt, sondern die Praxis des Programms miteinbezieht. Die explizierten Programmtheorien bilden die Grundlage für die Konzipierung, das Design und die Durchführung der Evaluation; sie werden in der Datenauswertung und -interpretation genutzt und bilden damit die Basis für die Bewertungen, die jede Evaluation vorzunehmen hat (vgl. Coryn et al. 2011, S. 201; Weiss 1995).

Mögliche Quellen, die Evaluatorinnen und Evaluatoren zum Erschließen von Programmtheorien nutzen können, sind u.a. Leitlinien und schriftlich niedergelegte Konzepte von Programmen, Projekten und Maßnahmen. Da Programmdurchführende auf der Grundlage ihrer persönlichen Erfahrungen und ihrer fachlichen Annahmen darüber handeln, wie ihre Aktivitäten Veränderungen oder

Stabilisierungen herbeiführen, ist es notwendig, auch diese Erfahrungen und Annahmen herauszuarbeiten. Mit Hilfe von Beobachtungen und Befragungen können solche vorwiegend impliziten Programmtheorien rekonstruiert werden (vgl. z.B. Haubrich 2009; Giel 2015).

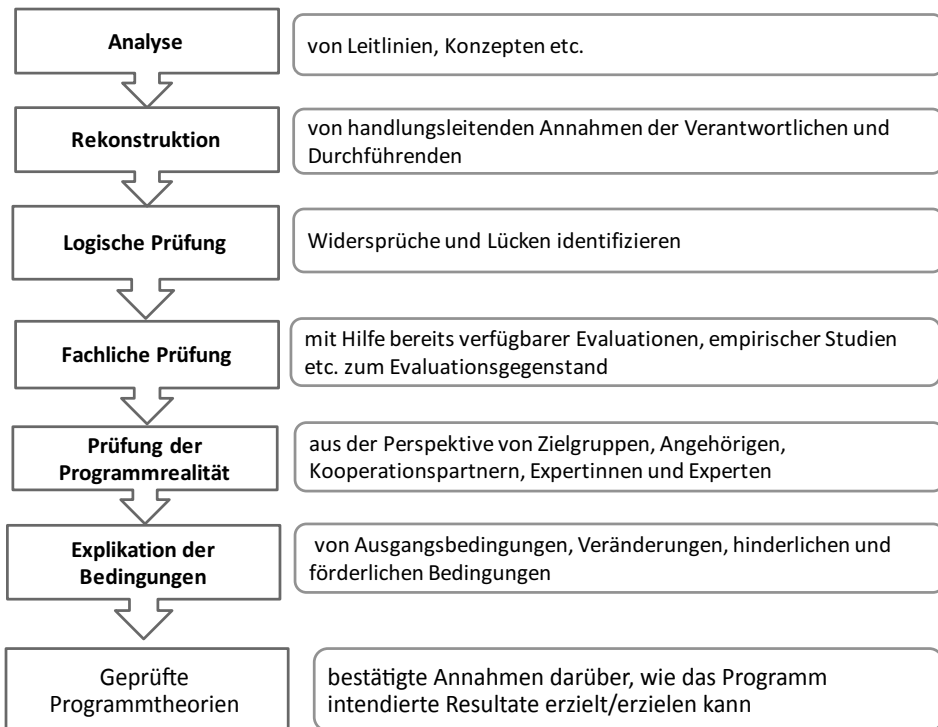
Natürlich besteht die Evaluation nicht nur darin, die das Programm strukturierenden und konturierenden Annahmen sichtbar zu machen, sondern diese auch zu prüfen. Zunächst ist eine logische Prüfung durchzuführen, die Widersprüche und vor allem Lücken identifiziert. Eine weitere Überprüfung aus externer Sicht erfolgt durch einen Abgleich mit zugänglichen, für das Programm relevanten Forschungs- und Evaluationsergebnissen. Schließlich gilt es, die Annahmen aus dem Programm und über das Programm mit der Wirklichkeit des Programms abzugleichen. Dabei wird etwa folgenden Fragen nachgegangen: Wird das Programm umgesetzt wie geplant? Werden mit den umgesetzten Aktivitäten die intendierten Resultate erzielt? Unter welchen Bedingungen lassen sich mit spezifischen Interventionen die intendierten Resultate erreichen, unter welchen Bedingungen ist das nicht der Fall?

Hierbei steht Evaluatorinnen und Evaluatoren das gesamte Repertoire an Methoden der Daten- und Informationsbeschaffung zur Verfügung: Befragungen, Beobachtungen, Tests, Sekundärdatenanalysen. Ob die Daten in eher standardisierter oder offener Weise erhoben werden, ist dabei von den Gegebenheiten im Programm, dem Auftrag an die Evaluation und nicht zuletzt von den Programmtheorien abhängig. Einen Überblick über die Bestandteile einer auf Programmtheorien basierenden Evaluation gibt Abbildung 1 auf S. 155.

Auch wenn dieses Schema eine stringente Abfolge suggeriert, so ist in der Evaluationspraxis, gerade bei Programmen, die sich noch in der Entwicklung befinden, von einer flexiblen Umsetzung auszugehen. Eine strikte Unterteilung in zunächst die Explizierung bzw. Rekonstruktion der Programmtheorie, dann deren Überprüfung ist dem Evaluationsgegenstand nicht angemessen. Wenn Programme noch experimentieren und Suchbewegungen nach passenden Antworten unternehmen, dann empfiehlt es sich, die Rekonstruktion der Programmtheorie mit deren Überprüfung zu verschränken. Damit bleibt Spielraum für Programmverbesserungen, -anpassungen und -weiterentwicklungen – bereits im Verlauf der Evaluation.

Ein Vorgehen, bei dem Erhebungsergebnisse mit den Programmbeteiligten kontinuierlich geteilt werden, ermöglicht nicht nur einen hohen Nutzen für das Programm selbst. Evaluatorinnen und Evaluatoren erhalten in diesem Rahmen immer wieder neue Informationen über die tatsächliche Umsetzung und die damit gemachten Erfahrungen. Diese Daten werden dann systematisch für die Weiterentwicklung und Überprüfung der Programmtheorie genutzt.

Abb. 1: Bestandteile einer Evaluation, die auf Programmtheorien basiert



Quelle: eigene Darstellung

Zusammenfassend lässt sich das hier umrissene Evaluationskonzept dadurch charakterisieren, dass Programmtheorien den Dreh- und Angelpunkt von Evaluation bilden. Sie rahmen den gesamten Prozess und bestimmen, welche Aspekte empirisch zu untersuchen sind. Dabei gibt das Programm den Takt vor und nicht von außen an das Programm herangetragene Theorien oder ein wenig flexibles Untersuchungsdesign, wie der zu Beginn genannte randomisierte Kontrollgruppenvergleich.

4. Programmtheorien bei Evaluationen am Beispiel von „Willkommensklassen“

Die im vorangegangenen Abschnitt aufgezeigten idealtypischen Schritte sollen im Folgenden an einem hypothetischen Beispiel illustriert werden. Im Mittelpunkt stehen hierbei:

- mögliche Fragestellungen von auf Programmtheorien basierenden Evaluationen im Kontext Schule,
- mögliche Quellen für Programmtheorien,

- Möglichkeiten zur Überprüfung der Programmtheorien anhand der Programmwirklichkeit und
- Strategien zur Bewertung von Programmen entlang der Programmtheorie.

Die Wahl für das Beispiel fällt auf die Integration von geflüchteten Kindern in die Regelschulen. In Berlin werden hierzu „Lerngruppen für Neuzugänge ohne Deutschkenntnisse“ eingerichtet, die auch „Willkommensklassen“ genannt werden; in Bayern spricht man von „Übergangs-“ und in Nordrhein-Westfalen von „Seiteneinsteigerklassen“. Angenommen wird eine Evaluation in Berlin, die den Auftrag erhält, Aufschluss darüber zu geben, wie das Lernangebot zu gestalten ist, damit die Integration in die Regelschule gut gelingt.

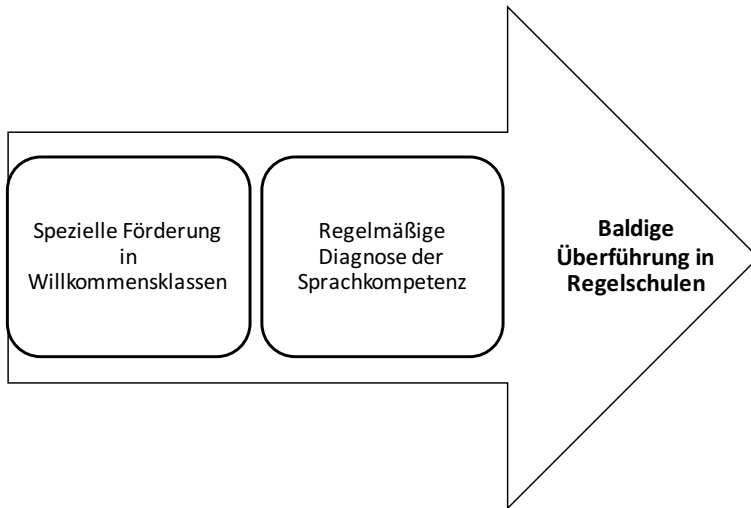
Zu berücksichtigen ist, dass Handlungsdruck besteht und dass Entscheidungen auf der Basis von wenig gesicherten Informationen getroffen werden müssen. In dieser Situation ist eine Evaluation, die eine gute Praxis identifizieren und beschreiben kann, um Entscheiderinnen und Entscheider sowie Lehrkräfte mit nützlichen Informationen und Anregungen für eine gelingende Umsetzung zu versorgen, angemessen. Diesen Bedarf deckt eine Evaluation, die Programmtheorien in den Mittelpunkt stellt, indem sie den Fokus darauf legt, *wie* überhaupt die Vermittlung von sprachlichen Kompetenzen gelingen bzw. *wie* den geflüchteten Kindern die Integration in Klassen der Regelschule ermöglicht werden kann.

Eine *erste Analyse offizieller Dokumente* untersucht die Grundannahmen des Programms. Hinweise liefert der Tagesspiegel vom 20. Juni 2015, der von einem Interview mit dem Berliner Bildungsstaatssekretär berichtet:

„Ziel des Unterrichts ist die schnellstmögliche Integration der Kinder und Jugendlichen in den regulären Unterricht.‘ Damit das klappt, sollen die Lehrer alle paar Monate die Deutschkenntnisse der Schüler prüfen und sie je nach Wissensstand, spätestens aber nach einem Schuljahr, in die reguläre Klasse verabschieden.“ (URL: <http://www.tagesspiegel.de/berlin/fluechtlings-kinder-in-der-schule-willkommensklassen-zwischen-integration-und-abschiebung/11910496.html>; Zugriffsdatum: 11.04.2016)

Im Einklang mit dieser öffentlichen Aussage steht der zentrale Handlungsleitfaden für Lehrkräfte „Von der Lerngruppe für Neuzugänge ohne Deutschkenntnisse in die Regelklasse. Ein dokumentierendes Verfahren“, das vom Landesinstitut für Schule und Medien Berlin-Brandenburg (LISUM 2014) herausgegeben wird. In diesem Leitfaden liegt der zentrale Akzent auf verschiedenen Diagnoseverfahren, Fragen- und Kriterienkatalogen zur Feststellung der deutschen Sprachkompetenz von Kindern. Nach dieser – hier natürlich nur verkürzt dargestellten – Sekundärdatenanalyse könnte die Evaluation mit einem ersten Entwurf für eine noch sehr grobe Programmtheorie starten (siehe Abbildung 2).

Abb. 2: Beispiel für eine erste grobe Programmtheorie



Quelle: eigene Darstellung

Es ist davon auszugehen, dass die den „Willkommensklassen“ unterliegenden Programmtheorien wesentlich ausdifferenzierter sind. Bereits eine erste *logische Prüfung* deckt auf, dass der Zeitraum der Separierung der geflüchteten Kinder von Schülerinnen und Schülern der Regelschulen möglichst kurz gehalten werden soll, um möglichst frühzeitig die Integration zu starten. Die zentrale Intervention wird in unterschiedlichen Test- und Diagnoseverfahren gesehen. Damit soll gewährleistet werden, dass kein Kind zu lange in den Willkommensklassen bleibt. Jedoch wird nicht deutlich, wie die eigentliche Vermittlung von deutscher Sprachkompetenz geschehen soll. Hier ist eindeutig eine Lücke in der Programmtheorie feststellbar.

In der Praxis wird es dazu sicherlich Konzepte und Strategien geben. Zur weiteren Differenzierung der Programmtheorie müsste die Evaluation in einem nächsten Schritt *rekonstruieren*, wie in den Willkommensklassen gearbeitet wird, konzeptionell wie auch ganz konkret und praktisch. Neben einer Recherche in den Konzeptpapieren der Träger der Willkommensklassen und in den Fortbildungsangeboten des Berlin-Brandenburger Lehrerfortbildungsinstituts gilt es vor allem herauszufinden, welche Lehrmaterialien und -medien, welche Lehr- und Lernmethoden umgesetzt werden. Hier bieten sich Einzel- und Gruppenbefragungen von Lehrerinnen und Lehrern sowie teilnehmende Beobachtungen von Unterrichtssequenzen an. Es ist davon auszugehen, dass diese Daten weitere Erkenntnisse über die Aktivitäten und Wirkannahmen der Programmumsetzenden liefern, etwa dazu, *wie die Bildung der Lerngruppen organisatorisch verläuft, welche Unterrichtsmethoden eingesetzt werden, welche Lernziele damit erreicht werden sollen und wie der Übergang in die Regelschulen gestaltet wird.*

Diese Wirkannahmen gilt es in einem nächsten Schritt einer *fachlichen Prüfung* zu unterziehen. Hierbei bietet sich ein Rückgriff auf konstruktivistische Lerntheorien an. Aus diesen weithin akzeptierten theoretischen Grundannahmen über erfolgreiches Lernen lassen sich Prüffragen an die Wirkannahmen der Durchführenden richten: Inwieweit enthalten die eingesetzten Methoden, Materialien und Medien aktivierende Elemente, wie breit sind die Angebote an Lernmöglichkeiten? Wodurch wird der Zugang zu den Lernangeboten ermöglicht? Ausführungen hierzu müssen an dieser Stelle notwendigerweise allgemein bleiben; für eine Konkretisierung und Präzisierung wäre eine tatsächlich durchgeführte Evaluation notwendig. Zur groben Orientierung lassen sich zunächst logisch und fachlich geprüfte Annahmen darüber ableiten, wie durch die Willkommensklassen der Übergang in Regelschulen gelingen soll (vgl. Abbildung 3).

Abb. 3: Programmtheorie nach logischer und fachlicher Prüfung



Quelle: eigene Darstellung

Bislang wurden die Programmtheorien, also die Annahmen darüber, wie mit den Willkommensklassen eine Integration der geflüchteten Kinder gelingen *kann* und *soll*, aus Programmsicht und aus fachlicher, durchführender Perspektive herausgearbeitet und expliziert. Inwieweit dies in der Umsetzung gelingt, das muss nun an der *Programmrealität überprüft* werden. Neben allgemeinen Daten zur Programmumsetzung (z.B. wie viele/welche Lehrkräfte im Einsatz sind, wie viele/welche Kinder an den Willkommensklassen teilnehmen, wie viele/welche Kinder in Regelschulen wechseln) sind die Annahmen unbedingt auch aus Zielgruppensicht zu überprüfen.

Die Überprüfung aus Zielgruppensicht stellt darauf ab, wie die Lernangebote von den geflüchteten Kindern aufgenommen werden, wie zielgruppenadäquat die Lernarrangements gestaltet sind, wie die Lernangebote von den Kindern genutzt werden. Daneben ist die Einschätzung der Kinder dazu wichtig, inwieweit sie sich auf den Übergang vorbereitet fühlen, bzw. das Urteil derjenigen, die bereits in den Regelschulen angekommen sind, wie sie den Übergang erlebt haben. Wahrscheinlich bietet es sich an, auch die Einschätzungen der Familienangehörigen bzw. des sozialen Umfelds wie auch die der Lehrerinnen und Lehrer der aufnehmenden Regelschulklassen zum Lern- und Integrationserfolg einzubeziehen. Zur Feststellung der Lernerfolge können darüber hinaus die dokumentierten Diagnoseverfahren genutzt werden.

Ein Abgleich zwischen den rekonstruierten Programmtheorien und den Daten der Überprüfung der Programmrealität aus den diversen Perspektiven gibt Aufschluss darüber, inwieweit die Annahmen als zutreffend eingeschätzt werden können. Anders ausgedrückt lassen sich bewertende Aussagen dazu ableiten, ob das pädagogische Handeln die gewünschten Resultate tatsächlich erzielt. Unter der Aufgabenstellung, gute Praxis zu identifizieren, ist es sicherlich auch notwendig, die *Bedingungen*, unter denen eine Integration (entlang der Programmtheorie) gelungen bzw. nicht gelungen ist, *auszuleuchten*. Folgende Aspekte sind neben anderen von Bedeutung:

- die Zusammensetzung der Zielgruppe, bspw. hinsichtlich Alter und Altersspanne in der Gruppe, Bildungsstand, besonderer Belastungen, der Fluktuation in der Gruppe etc.;
- zur Verfügung stehende Ressourcen: Qualifikationen der umsetzenden Lehrkräfte, Ausstattung der Räume, Ausstattung mit Lehr- und Lernmaterial;
- Kooperationen zwischen Willkommensklassen, Regelschulbetrieben und weiteren unterstützenden Angeboten.

Am Ende steht die Beschreibung gelingender Praxis, also eine Darstellung, welche konkreten pädagogischen Interventionen unter welchen ausgewiesenen Bedingungen intendierte Resultate erzielen können.

In diesem konstruierten Beispiel steckt so viel Realität, dass absehbar ist, dass in der Programmumsetzung experimentiert wird, dass viele situative Entscheidungen getroffen werden und weniger auf langfristig entwickelte Konzepte zurückgegriffen werden kann. Ebenso ist davon auszugehen, dass sich die Praxis permanent verändert, die eingesetzten Fachkräfte Erfahrungen sammeln, daraus Schlüsse ziehen und sich den veränderlichen Bedingungen anpassen. Evaluatorinnen und Evaluatoren gehen diese Bewegungen mit und nehmen kontinuierlich Bewertungen vor: Sie bewerten, ob die Programmtheorien logisch und fachlich konsistent sind, ob sich die Annahmen auch in der Praxis bewähren, und finden heraus, welche Annahmen der Programmdurchführenden in der praktischen Umsetzung scheitern oder nicht die intendierten Resultate zeigen. Daraus lassen sich jeweils Optimierungsvorschläge ableiten. Die

Evaluierenden stellen durch regelmäßigen Austausch mit Programmverantwortlichen und -umsetzenden sicher, dass die Programmtheorien auf dem aktuellen Stand sind, und die Programmakteure können kontinuierlich ihr Handeln datenbasiert verbessern.

5. Reflexion der Nutzungspotenziale von Evaluationen, die sich auf Programmtheorien stützen

Der vorangestellte Entwurf für die Umsetzung einer auf Programmtheorien basierenden Evaluation muss notwendigerweise – da es sich um ein hypothetisches Beispiel handelt – oberflächlich bleiben. Trotzdem sollte deutlich geworden sein, dass sich mit dem hier präsentierten Konzept Evaluationen deutlich von Lernstandsmessungen und in Schul-Selbstevaluationstools verbreiteten Selbsteinschätzungen unterscheiden. Das wichtigste Kennzeichen theoriebasierter Evaluationen besteht darin, dass sie nicht lediglich Resultate deskriptiv erfassen, sondern konsequent den Zusammenhang zwischen (pädagogischen) Aktivitäten und deren Resultaten untersuchen. Damit steht durchgängig die Wirkfähigkeit von Programmen auf dem Prüfstand. Ein solches Vorgehen schafft Transparenz und trägt erfahrungsgemäß dazu bei, das Wirkungspotenzial von Programmen realistischer einzuschätzen.

Die sukzessive Entwicklung und Überprüfung der Programmtheorien ermöglicht es, mit den Entwicklungen von solchen Programmen mitzugehen, die sich selbst noch auf der Suche nach Lösungen befinden, Neues ausprobieren und Umsetzungsstrategien revidieren. Eine auf Programmtheorien beruhende Evaluation ist in der Lage, diese Suchbewegungen und Entwicklungen nachzuvollziehen und sich flexibel den Bedingungen im Programm anzupassen.

Zusätzlich unterstützt die Evaluation die Entwicklung des Programms. Sie ermöglicht es, genauer zu verstehen, wieso Erfolge eintreten oder ausbleiben. Dadurch, dass die Programmtheorien nicht schlicht von Evaluierenden gesetzt, sondern im kontinuierlichen Austausch mit Programmakteuren, Beteiligten und Betroffenen entwickelt und mit externen theoretischen Wissensbeständen angereichert werden, ist ein fortwährendes Lernen abgesichert. Durch die Explikation förderlicher und womöglich auch hinderlicher Bedingungen ist sichergestellt, dass mögliche Stellschrauben identifiziert werden. Die Berücksichtigung dieser Bedingungen ermöglicht es, transferfähiges Wissen zu generieren und dieses Wissen an andere Standorte und in die Fachöffentlichkeit weiterzugeben. Unter Einbeziehung der Praxis erarbeitete Programmtheorien gewährleisten, dass sich Theorie und Praxis gegenseitig befruchten.

Das hypothetische Beispiel hat deutlich gemacht, dass eine auf Programmtheorien basierende Evaluation durchaus aufwändig und anspruchsvoll ist und somit von Schulen in Eigenverantwortung – im Sinne von individueller, kollegialer Selbstevaluation – nicht zu „stemmen“ ist und dass nur wenige Schulen über die erforderlichen Kompetenzen für solche anspruchsvollen Inhouse-Evaluationen verfügen dürften (vgl. Beywl/Balzer in diesem Heft, S. 191-204). Jedoch kann der konsequente Rückbezug auf die Frage nach dem „Wie“ eine wertvolle Anregung für die verschiedenen Arten schulinterner Evaluationen liefern. Der Fokus darauf, wie und mit welchen Interventionen, auf der Grundlage welcher theoretischer Wirkannahmen Lehrerinnen und Lehrer zum Lernerfolg beitragen wollen, befördert wirkungsorientiertes Handeln – und auch Evaluieren.

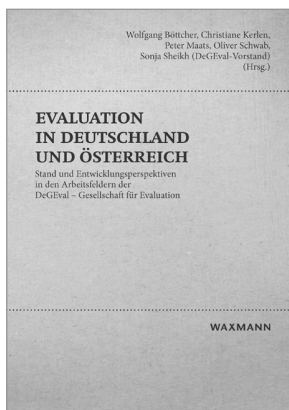
Literatur und Internetquelle

- Beywl, W./Balzer, L. (2016): Aufbau von Evaluationskompetenzen für interne Schulevaluation durch projektbezogene Fortbildung. In: Die Deutsche Schule 108, H. 2, S. 191-204.
- Bickman, L. (1987): The Function of Program Theory. In: Ders. (Hrsg.): Using Program Theory in Evaluation. New Directions for Program Evaluation. San Francisco, CA: Jossey-Bass, S. 5-18.
- Böttcher, W./Hense, J. (2016): Evaluation im Bildungswesen – eine nicht ganz erfolgreiche Erfolgsgeschichte. In: Die Deutsche Schule 108, H. 2, S. 117-135.
- Coryn, C.L./Noakes, L.A./Westine, C.D./Schröter, D.C. (2011): A Systematic Review of Theory-driven Evaluation Practice from 1990 to 2009. In: American Evaluation Association (Hrsg.): American Journal of Evaluation 32, H. 2, S. 199-226.
- Giel, S. (2013): Theoriebasierte Evaluation. Konzepte und methodische Umsetzungen. Münster u.a.: Waxmann.
- Giel, S. (2015): Wirkungen auf der Spur. In: Giel, S./Klockgether, K./Mäder, S.: Evaluationspraxis. Professionalisierung – Ansätze – Methoden. Münster u.a.: Waxmann. S. 111-130.
- Griese, C. (2016): Evaluation in der Schule. In: Griese, C./Marburger, H./Müller, T. (Hrsg.): Bildungs- und Bildungsorganisationsevaluation. Ein Lehrbuch. Berlin: de Gruyter, S. 163-187.
- Haubrich, K. (2009): Sozialpolitische Innovation ermöglichen. Die Entwicklung der rekonstruktiven Programmtheorie-Evaluation am Beispiel der Modellförderung in der Kinder- und Jugendhilfe. Evaluation innovativer multizentrischer Programme. Münster u.a.: Waxmann.
- Kromrey, H. (2001): Evaluation – ein vielschichtiges Konzept. In: Sozialwissenschaften und Berufspraxis 24, H. 2, S. 105-131.
- Kromrey, H. (¹²2009): Empirische Sozialforschung. Überarb. und ergänzte Aufl. Stuttgart: Lucius & Lucius.
- LISUM (Hrsg.) (2014): Von der Lerngruppe für Neuzugänge ohne Deutschkenntnisse in die Regelklasse. Ein dokumentierendes Verfahren. URL: http://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/unterricht/unterrichtsentwicklung/Durchgaengige_Sprachbildung/Publicationen_sprachbildung/Lerngruppe_fuer_Neuzugaenge_ges_WEB_2014_12.pdf; Zugriffsdatum: 30.12.2015.

- Pawson, R./Tilley, N. (2004): Realistic Evaluation. London/Thousand Oaks, CA/New Delhi: Sage.
- Rogers, P.J. (2000): Program Theory. Not Whether Programs Work, but How They Work. In: Stufflebeam, D.L./Madaus, G.F./Kellaghan, T. (Hrsg.): Evaluation Models. Viewpoints on Educational and Human Services Evaluation. Boston, MA/Dordrecht/London: Kluwer Academic Publishers, S. 209-232.
- Weiss, C.H. (1995): Nothing As Practical As Good Theory: Exploring Theory-based Evaluation for Comprehensive Community Initiatives. In: Connell, J.P./Kubisch, A.C./Schorr, L.B./Weiss, C.H.: New Approaches to Evaluation Community Initiatives. Concepts, Methods, and Contexts. Washington: Aspen Institute.

Susanne Giel, Dr., geb. 1963, Gesellschafterin und Projektleiterin bei Univation – Institut für Evaluation GmbH.
Anschrift: Univation – Büro Berlin, Gubener Straße 25, 10243 Berlin
E-Mail: susanne.giel@univation.org

UNSERE BUCHEMPFEHLUNG



2014, 220 Seiten, br., 29,90 €,
ISBN 978-3-8309-3149-2
E-Book: 26,99 €,
ISBN 978-3-8309-8149-7

Wolfgang Böttcher, Christiane Kerlen,
Peter Maats, Oliver Schwab, Sonja Sheikh
(DeGEval-Vorstand) (Hrsg.)

Evaluation in Deutschland und Österreich

Stand und Entwicklungsperspektiven
in den Arbeitsfeldern der DeGEval

Mit diesem Band stellt die DeGEval – Gesellschaft für Evaluation die Ergebnisse ihrer Tätigkeit in verschiedenen Evaluationsfeldern in Deutschland und Österreich vor. Die Berichte gewähren nicht nur einen Einblick in die Praxis und historische Entwicklung der Evaluation in verschiedenen Politikfeldern unter Betrachtung der jeweils spezifischen Anforderungen und Problemstellungen, sondern liefern zugleich einen anschaulichen Überblick über den aktuellen Stand der Evaluationsforschung.



www.waxmann.com