

Claudia Bergmüller

# Wirkungen beobachten? Anregungen schulischer Evaluationsforschung für Globales Lernen im (außer-)schulischen Kontext<sup>1</sup>

## Zusammenfassung:

Im Zuge immer knapper werdender öffentlicher Finanzmittel muss sich auch der Bereich der entwicklungsbezogenen Inlandsarbeit zunehmend der Forderung nach einer stärkeren Wirkungsorientierung der pädagogischen Arbeit stellen. Diese Forderung gilt sowohl für eine wirkungsorientierte Planung von Bildungsmaßnahmen als auch für eine an Wirkungsanalysen orientierte Berichterlegung hinsichtlich des Erfolgs dieser Maßnahmen. Die im Rahmen dieser Analysen nötige Beobachtung von Wirkungen ist im Kontext Globalen Lernens allerdings nicht trivial. In diesem Beitrag wird dieser Umstand als Ausgangspunkt genommen, um Möglichkeiten nachzugehen, Wirkungsbeobachtungen im Globalen Lernen vor allem auch unter der Perspektive praxisbezogener Berichterlegung angemessen umzusetzen. Ausgehend von zentralen Herausforderungen von Wirkungsbeobachtungen im Globalen Lernen wie zum Beispiel dem Technologiedefizit pädagogischer Arbeit und der damit verbundenen Unterkomplexität von Wirkungsketten werden Bezug nehmend auf den Arbeitsbereich der schulischen Evaluationsforschung Anregungen gegeben, wie mit diesen Herausforderungen konstruktiv umgegangen werden kann.

**Schlüsselworte:** *Globales Lernen, Wirkungsbeobachtung, Evaluation, Evaluationsforschung*

## Abstract:

In the wake of increasingly scarce public funds also the area of development education is increasingly calling for a more powerful outcome resp. impact orientation of the educational work. This requirement applies to both a pro-active planning of educational activities as well as a report writing based on impact analysis evaluating the success of these measures. However, the observation of effects which is necessary for such an analysis is not trivial in the context of global education. In this paper, this fact is taken as a starting point to investigate possibilities to appropriately implement effect observation in the field of Global Education especially from the perspective of a praxis-related reporting. Based on experiences of school evaluation research suggestions will be made how to deal constructively with key challenges as the technology deficit of educational work and the associated under-complexity of causal resp. functional chains.

**Keywords:** *global education, effect observation, evaluation, evaluation research*

## Einleitung

In nahezu allen Steuerungsbereichen des öffentlichen Lebens vollzieht sich augenblicklich ein Paradigmenwechsel des Steuerungswesens weg von der Input- hin zur Output-Steuerung. Dahinter steht die Vorstellung, Entwicklungen in den jeweiligen Bereichen nicht mehr nur über die Inputseite – und damit z.B. über Regeln, Vorschriften, Bestimmungen etc. – zu initiieren und zu lenken, sondern stärker vom Output – von Vorgaben zu Ergebnissen, Leistungen und Wirkungen – her. Mit der Überprüfung der erreichten Qualität dieses Output verbindet sich zum einen die Erwartung, Aufschluss darüber zu erhalten, an welchen Stellen Inputs oder mit den Inputs verbundene Prozesse optimiert werden können oder müssen, um die Outputqualität zu halten oder zu steigern. Zum anderen soll – vor allem durch die immer knapper werdenden öffentlichen Finanzmittel – durch die Überprüfung der Outputqualität Rechenschaft über die angemessene Verwendung eingesetzter Mittel abgelegt bzw. der jeweilige Mitteleinsatz entsprechend legitimiert werden.

Auch in der Bildungsarbeit von Nichtregierungsorganisationen (NROs) spiegelt sich diese Outputorientierung wider: Die ursprünglich im Kontext der Verpflichtung zu gemeinsamen Entwicklungszielen im Rahmen der „Millenniums-Erklärung“ und der „Paris Deklaration“ erhobene Forderung an Gebergemeinschaft und NROs, glaubwürdig nachzuweisen, dass sie die ihnen zur Verfügung gestellten Mittel effektiv einsetzen und einen sichtbaren Beitrag zur Entwicklung in den Partnerländern leisten“ (Deutsche Welthungerhilfe e.V. 2008, S. 1) bezieht sich über die Arbeit in den Partnerländern hinaus nun auch auf die Inlandsarbeit dieser Akteure. Dieser Trend erhält nicht zuletzt dadurch Rückenwind, dass vor allem in denjenigen Fällen, in denen Akteure der Inlandsarbeit gegenüber Fördergebern Rechenschaft über ihre Mittelverwendung ablegen müssen, Ist-Soll-Vergleiche Konjunktur zu haben scheinen (vgl. u.a. Engagement Global/FEB 2011; Stiftung Nord-Süd-Brücken 2011).

Diese Entwicklung hat dazu geführt, dass stark darum gerungen wird, was in der entwicklungsbezogenen Bildungsarbeit wie als (erreichbare) Wirkung beschrieben und beobachtet werden kann. In diesem Beitrag werden aus der Perspektive

schulischer Evaluationsforschung Anregungen für diese Debatte gegeben. Dabei werden zunächst einige Herausforderungen beschrieben, die sich bei der Beobachtung und Analyse von Wirkungen im Globalen Lernen stellen. Danach wird mit Blick auf Erfahrungen aus der schulischen Evaluationsforschung der Frage nachgegangen, wie mit diesen Herausforderungen umgegangen werden kann.

### **Herausforderungen von Wirkungsbeobachtungen im Globalen Lernen**

#### **Das Technologiedefizit von Bildungsarbeit**

Die Beobachtung und Analyse von Wirkung zielt darauf ab, Veränderungen zu erfassen, die durch bestimmte Maßnahmen ausgelöst wurden. Dabei gilt es zum einen, die Veränderungen selbst zu erheben, und zum anderen die einzelnen Faktoren, die zu diesen Veränderungen geführt haben, zu ermitteln und entsprechend zu gewichten, um auf diese Weise den Beitrag der jeweiligen Maßnahme zu den Veränderungen analysieren zu können (vgl. hierzu u.a. Seitz 2010, S. 14; VENRO 2010, S. 1).

Im Globalen Lernen ist die Analyse derartiger Wirkungszusammenhänge allerdings nicht trivial. Im Positionspapier 2/2010 zur Wirkungsorientierung konstatiert VENRO – wenngleich zunächst im Bezug auf Maßnahmen der Entwicklungszusammenarbeit – dass „jede Maßnahme [...] unterschiedliche Wirkungen haben [kann], die in ihrer Gesamtheit und als ein sich gegenseitig beeinflussendes Wirkungsgeflecht in ihrer Komplexität oft nicht vollständig erfasst und beschrieben werden können“ (VENRO 2010, S. 1). Für Bildungsmaßnahmen im Allgemeinen und Maßnahmen Globalen Lernens im Besonderen gilt dies in verstärktem Maße. Mit Maßnahmen Globalen Lernens sollen Menschen motiviert und befähigt werden, an der Gestaltung der Weltgesellschaft engagiert und sachkundig teilzuhaben (vgl. BMZ 2007). In vielen Konzepten wird daher als Ziel formuliert, die Teilnehmenden für Globalisierungsfragen zu sensibilisieren, vernetztes Denken zu fördern, die Mitwirkung des Einzelnen und dessen Verantwortungsbewusstsein im Hinblick auf die Bewältigung globaler Herausforderungen in Bereichen wie Entwicklung, Umwelt, Frieden und Menschenrechte anzuregen sowie dazu anzuhalten, den eigenen Lebensstil vor dem Hintergrund sozialer bzw. ökologischer Folgen zu überdenken. Über den Aufbau von Wissen hinaus geht es somit auch um den Aufbau bestimmter Einstellungen und Werthaltungen, sozialer Kompetenzen sowie einer „reflektierten Handlungsfähigkeit“ (Jungk 2010, S. 19).

Entgegen vieler Maßnahmen in der Entwicklungszusammenarbeit (EZ), die – vor allem im Rahmen von technischer EZ – oftmals einen stärkeren Kausalzusammenhang aufweisen (vgl. u.a. das Beispiel Brunnenbau in Deutsche Welthungerhilfe e.V. 2008, S. 5), ist das hier intendierte Lernen bzw. der hier intendierte Bildungsprozess der Beteiligten ein sehr komplexer Vorgang,

- der zeitlich versetzt von Bildungsangeboten sein kann,
- in dem sich Wissen nicht zwingend auf Verhalten auswirken wird und
- der nach Luhmann & Schorr (1979, S. 345) selbstreferentiell ist, also dem subjektiven Selbstbewusstsein, Denken und Wollen unterliegt, und daher von außen nicht kausal stringent auf bestimmte Außenreize zurückzuführen ist.

In Bildungszusammenhängen besteht diese Selbstreferentialität sogar in doppelter Hinsicht, da hier mindestens zwei Subjekte (die Lehrperson und der/die Lernende) über die Bildungsarbeit in eine Beziehung zueinander gesetzt werden und damit (mindestens) zwei Selbstreferenzen aufeinander treffen. Will man Bildungsarbeit als kausalen Zusammenhang beschreiben, müsste für jede Wirkung, die durch diese Beziehung bezweckt wird, dieser Zusammenhang so beschrieben werden, dass die selbstreferentiellen Prozesse (also z.B. das Denken oder Wollen) jedes Beteiligten mitantizipiert werden. Dies gestaltet sich bei einer Beziehung zwischen zwei Subjekten bereits äußerst schwierig, bei Gruppen ist dies geradezu unmöglich. Dieses von Luhmann & Schorr als ‚Technologiedefizit‘ (vgl. ebd.) von Bildungsarbeit beschriebene Problem macht die Annahme von Kausalzusammenhängen damit schwierig und die Zurechenbarkeit von Bildungsmaßnahme und Bildungsergebnis schwer bestimmbar (vgl. ausführlich Scheunpflug 2004; Bergmüller & Scheunpflug im Erscheinen).

Zudem ist Bewusstsein nicht beobachtbar, so dass die im Globalen Lernen angestrebten Bewusstseins- und Einstellungsveränderungen nach Seitz (2010, S. 15; vgl. ähnlich auch Jungk 2010, S. 19) von außen nur schwer zugänglich sind und sich nur stellvertretende Indikatoren wie z.B. bewertende Äußerungen, abfragbares Wissen oder in bestimmten Situationen gezeigte Verhaltensweisen angeben lassen, die aber wiederum diesen Bewusstseins- und Einstellungsveränderungen nicht unbedingt unmittelbar zurechenbar sein müssen (vgl. zur Darstellung dieses Zusammenhangs ausführlicher Seitz 2010, S. 15). Hinzu kommt die Schwierigkeit, dass sich auch die jeweilige Ausgangssituation, auf die die Wirkungsbeobachtung Bezug nimmt, nicht trivial bestimmen lässt: Noch relativ machbar scheint dies bei der Erhebung von Vorwissen oder thematischem Interesse. Schwieriger wird es bei der Erhebung von Vorerfahrungen oder Lernausgangsbedingungen. Für Wirkungsbeobachtungen im Globalen Lernen besteht also ein ernst zu nehmendes Zurechnungsproblem. Zudem gilt natürlich auch hier das Problem, das für die Wirksamkeit aller Lernangebote von Bedeutung ist: Die Rezeption des Bildungsangebots durch die lernenden Subjekte ist nicht unabhängig von deren kognitiver Grundausrüstung, deren sozioökonomischem Hintergrund und deren Lernerfahrungen. Kompetenzmessungen wie jene aus PISA erheben deshalb immer auch den sozioökonomischen Hintergrund und die Intelligenz mit, um die sich durch diese Faktoren ergebenden Verzerrungen systematisch ausschließen zu können (vgl. z.B. in diesem Kontext das Angebot-Nutzungsmodell; Helmke 2009 siehe auch die Ausführungen unten).

#### **Die Unterkomplexität von Wirkungsketten**

Trotz des eben beschriebenen ‚Technologiedefizits‘ von Bildungsarbeit im Allgemeinen und von Globalem Lernen im Besonderen wird in der entwicklungsbezogenen Bildungsarbeit augenblicklich – vor allem von Fördergeberseite – die Verwendung von Wirkungsketten wie „Input – Aktivitäten – Leistungen – Nutzung der Leistung – direkte Wirkung – indirekte Wirkung“ (vgl. u.a. Massing et al. 2010) oder „Input – Aktivität – Output – Outcome – Impact“ (vgl. OECD 2009 sowie Quiring in diesem Heft) propagiert. Die einzelnen Elemente der Kette markieren dabei die jeweiligen Anschlussstellen in

Projektverläufen, angefangen bei den zur Verfügungstehenden technischen, materiellen und personellen Ressourcen (Inputs) über die mit diesen Ressourcen durchgeführten Aktivitäten zur Erstellung/Erbringung von Leistungen bis hin zu entsprechenden unmittelbaren oder mittelbaren Wirkungen, die durch die Leistungen angestoßen werden (vgl. Massing et al. 2010, S. 38).

Die durch Wirkungsketten suggerierte Möglichkeit, Programmzusammenhänge in komplexitätsreduzierter Form abbilden und damit handhabbar machen zu können, führt unter anderem zu der Erwartung, Antragsstellende sowohl zu einer realistischen Selbsteinschätzung ihrer pädagogischen Möglichkeiten anzuhalten und damit die Projektplanungen zu verbessern (vgl. u.a. Rosen 2010) als auch die auf diese Planung bezogene Berichts- bzw. Rechenschaftslegung nach Abschluss eines Projektes effizienter gestalten zu können (vgl. BMZ 2008). Für die Projektplanung mag sich diese Erwartung erfüllen. Für die Rechenschaftslegung allerdings entsteht hier ein Dilemma: Je mehr die Komplexität von Bildungsprozessen durch eine Darstellung in z.B. Kettenform reduziert wird, desto größer wird die Gefahr, Wirkungszusammenhänge zu unterkomplex abzubilden und Einflussfaktoren auszublenden, die aber beispielsweise für eine adäquate Einschätzung der Qualität eines Bildungsprojektes wesentlich wären. Dies sei in der Gegenüberstellung einer beispielhaften Wirkungskette (vgl. Abb. 1) und einem Wirkungsmodell aus der Schulforschung (Abb. 2) illustriert. Derartige so genannte Angebots-Nutzungs-Modelle (vgl. u.a. Helmke 2009) werden seit geraumer Zeit in der Unterrichtsforschung verwendet, um das Zustandekommen von Lernergebnissen zu erklären. In diesen Modellen wird ähnlich der o.g. Wirkungskette Unterricht als ein Angebot verstanden, welches von den Schülerinnen und Schülern in Form bestimmter Lernaktivitäten genutzt wird (vgl. hierzu auch Fend 1998). In diesen Modellen wird deutlich, dass das unterrichtliche Angebot nicht notwendiger Weise direkt zu den Wirkungen führt, sondern dass seine Wirksamkeit davon abhängt, (1) „ob und wie Erwartungen der Lehrkraft und unterrichtliche Maßnahmen von den Schülerinnen und Schülern überhaupt wahrgenommen und interpretiert werden sowie (2) ob und zu welchen motivationalen, emotionalen und volitionalen (auf den Willen

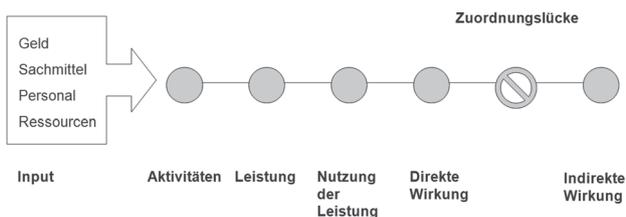


Abb. 1  
Quelle: Massing et al. 2010, S. 37

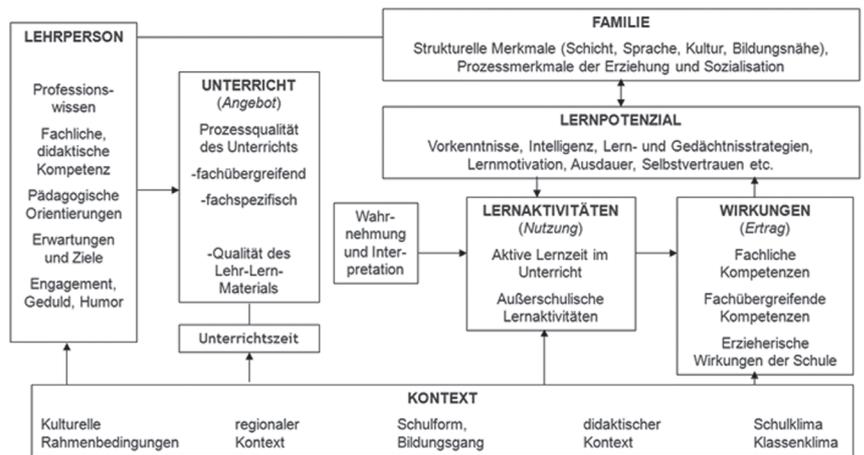


Abb. 2  
Quelle: Helmke 2009, S. 73

bezogenen) Prozessen sie auf Schülerseite führen“ (Helmke 2009, S. 74). Beides bedingt, ob, wie intensiv und wie nachhaltig die Schülerinnen und Schüler das unterrichtliche Angebot nutzen und beeinflusst somit wiederum die (möglichen) Wirkungen des Unterrichts.

### Die Festlegung des Beobachtungsgegenstandes

In der Frage, was als Wirkung beschrieben und beobachtet werden kann bzw. was sich auch gerade nicht (mehr) als ‚Wirkung‘ im Sinne einer auf eine bestimmte Maßnahme zurückzuführenden Veränderung beschreiben lässt, gibt es ein relativ großes Begriffsangebot, je nachdem, auf welches im Feld propagierte Wirkungsmodell man sich beziehen möchte oder muss. Quiring (in diesem Heft) beispielsweise geht in ihrem Beitrag von der Unterscheidung nach Output, Outcome und Impact aus (vgl. zur Unterscheidung von Output und Outcome alternativ das Accountability and Evaluation Glossary 2008). Massing et al. (2010) unterscheiden entsprechend der offiziellen BMZ- und DAC-Terminologie zwischen Nutzung der Leistung, direkten und indirekten Wirkungen (ebd. S. 39). Gohl (2002) thematisiert ergänzend zum Begriff ‚impact‘ den Begriff der ‚Nachhaltigkeit‘.

Egal welchen Begrifflichkeiten man sich anschließt, steht dahinter das Bemühen, verschiedene Wirkungsebenen voneinander zu unterscheiden und damit auch die Einflussmöglichkeiten des eigenen Handelns angemessen zu definieren. Zentral ist dabei die Unterscheidung, ob es sich bei den zu definierenden ‚Wirkungen‘ um unmittelbare, kurz- oder mittelfristige, direkt mit den Maßnahmenaktivitäten oder -leistungen in Zusammenhang stehende Veränderungen handelt, oder um längerfristige, indirekte Veränderungen, bei denen die Kausalbeziehung weniger klar ist, „entweder weil zwischen Projektmaßnahmen und Wirkungen mehrere Stufen sind, oder weil außer den Projektmaßnahmen noch mehrere andere Maßnahmen und Rahmenbedingungen bedeutenden Einfluss auf die Veränderungen gehabt haben können“ (Gohl 2002, S. 38). Grundsätzlich gilt: Je mittelbarer die Einflussmöglichkeiten, desto größer ist das bereits beschriebene Zurechnungsproblem und desto eher ist von indirekten Wirkungen zu sprechen.

Im Globalen Lernen werden Wirkungsbeobachtungen vor allem zu Evaluationszwecken eingesetzt. Theoretisch sollte

es somit – versteht man Evaluation dem eigentlichen Sinn nach als Entwicklungsinstrument – über das Ziel von Rechenschaftslegung hinaus darum gehen, mit Wirkungsbeobachtungen Erkenntnisse zu generieren, die dabei helfen, die Qualität der jeweiligen Bildungsarbeit zu verbessern. Deshalb wird in Evaluationshandreichungen angeraten, sich vor allem auf direkte Wirkungen zu konzentrieren, da hier eine zumindest einigermaßen gesicherte Zurechenbarkeit des eigenen Handelns auf erzielte Wirkungen gegeben ist.

In der Praxis hat – entsprechend der Ist-Soll-Logik von Verwendungsnachweisen – in vielen Fällen augenblicklich die Rechenschaftslegung den größeren Stellenwert. Deshalb sollen Wirkungsbeobachtungen häufig vor allem die Frage beantworten, inwiefern die mit einem Projekt intendierten Wirkungen auch tatsächlich erreicht wurden und diese Frage wird in vielen Bildungsmaßnahmen im Globalen Lernen als Frage danach formuliert, ob und inwiefern es gelungen ist, bestimmte Einstellungen und Haltungen anzubahnen und reflektiertes Handeln zu fördern. Damit zielen Wirkungsbeobachtungen aber eher auf indirekte Wirkungen ab, denn Bildung ist immer Selbstbildung (vgl. weiterführend Jungk 2010, S. 19) und zudem stark kontextabhängig (vgl. u.a. Helmke 2009).

Ganz abgesehen davon, dass die Formulierung von ‚Bewusstseins- und Einstellungsveränderungen‘ als Ziel pädagogischen Handelns angesichts des Beutelsbacher Konsens (vgl. u.a. Scherb 2010) und dem darin festgeschriebenen Verbot, Schülerinnen und Schülern auf vorgegebene Meinungen ‚einzuschwören‘, prinzipiell sehr vorsichtig vorgenommen werden sollte – ist es mit diesem Fokus auch schwierig, Ableitungen für eine Qualitätsverbesserung der eigenen Bildungsarbeit zu treffen, denn dies geht nur, wenn eine gewisse Zurechenbarkeit der durchgeführten Maßnahmen auf die erreichten Ziele/Wirkungen möglich ist.

#### **‚Wissenschaftliche‘ Anforderungen an Praxisarbeit – Die Diskrepanz zwischen Anspruch und Machbarkeit**

Die vierte hier zu thematisierende Herausforderung betrifft die Schwierigkeit, Qualitätsansprüche an Wirkungsbeobachtungen einerseits und die Limitierungen von Praxisarbeit andererseits in Balance zu bringen. Wirkungsbeobachtungen im Globalen Lernen sollen „kein wissenschaftliches Forschungsprojekt“ sein, heißt es in der Handreichung „Wirkt so!“ (Massing et al. 2010, S. 45). Sie müssen „innerhalb des üblichen Ressourcenrahmens eines Projektes durchführbar sein“ (ebd.) Dieser Ressourcenrahmen ist meist eng gesteckt und angesichts immer knapper werdender öffentlicher Finanzmittel ist auch nicht zu erwarten, dass sich diese Lage verbessern wird. Im Gegenteil. Aus Praxisperspektive ist eine solche Forderung deshalb nur allzu verständlich. Auf der anderen Seite wird aber unter anderem von Fördergeberseite – ebenso berechtigt – ein gewisser Qualitätsanspruch an Wirkungsbeobachtungen gefordert.

Wirkungsbeobachtungen, vor allem wenn sie von Antragstellenden im Rahmen ihrer Rechenschaftslegung durchgeführt werden, sind schon allein deshalb in den seltensten Fällen ‚wissenschaftliche Forschungsprojekte‘, weil sie sich ausschließlich auf ihr eigenes Projekt konzentrieren. Wissenschaft dagegen zielt darauf ab, über den Einzelfall hinausgehende Erkenntnisse zu generieren, die auch auf andere Bereiche anwendbar

sind. Daraus resultieren noch einmal andere Anforderungen an das methodische Vorgehen (vgl. zur Abgrenzung von Evaluation, Evaluationsforschung und Grundlagenforschung u.a. Bergmüller & Stadler-Altman 2009). Dennoch unterliegen Wirkungsbeobachtungen, egal ob sie von wissenschaftlicher Seite für Forschungszwecke oder von Praktikerinnen und Praktikern zur Rechenschaftslegung (und/oder Qualitätsentwicklung) durchgeführt werden, an folgenden Stellen vergleichbaren Qualitätsstandards: In beiden Fällen muss gewährleistet sein, dass bearbeitet wird, was zu bearbeiten vorgegeben wird, dass also reliable Daten erhoben werden und es für externe Personen nachvollziehbar ist, wie die Daten erhoben und ausgewertet wurden bzw. auf welcher Datengrundlage bestimmte Ableitungen gezogen werden (vgl. hierzu ausführlich Joint Committee on Standards for Educational Evaluation 2006; 2011). Diese Anforderungen sind für redliche Wirkungsbeobachtungen trotz aller Ressourcenengpässe bindend.

#### **Anregungen für das Handlungsfeld des Globalen Lernens**

Im Folgenden wird versucht, mit Blick auf die oben beschriebenen Herausforderungen ein paar Anregungen zu geben. Als Bezugsfeld dient dabei die schulische Evaluationsforschung. Gemeint sind damit Forschungsprojekte, die „eine bestimmte schulische Praxis in einem klar umrissenen Forschungsfeld daraufhin untersuchen, ob zuvor definierte oder im Prozess gewonnene Qualitätskriterien erreicht werden“ (vgl. Prengel et al. 2004, S. 185). Die Bedeutung dieses Forschungsbereichs hat – ebenso wie die Bedeutung von Evaluation im Globalen Lernen – in den letzten zehn bis 15 Jahren deutlich zugenommen. Vor allem seit PISA 2000 liegt ein verstärkter Fokus auf der systematischen Erforschung schulischer Rahmenbedingungen für das Lernen der Schülerinnen und Schüler mit dem Ziel, Steuerungswissen zu generieren, um Schul- und Unterrichtsqualität zu verbessern und das schulische Lernen und damit die Leistungsergebnisse zu optimieren. In zahlreichen Modellversuchen werden seither pädagogische Ansätze erprobt und ihre jeweiligen Wirkungen analysiert (vgl. zur Einführung in die schulische Evaluationsforschung Bergmüller & Stadler-Altman 2009; Prengel et al. 2004).

#### **Anregungen für die Festlegung des Beobachtungsgegenstandes**

In der schulischen Evaluationsforschung finden sich verschiedene Systematisierungen dessen, was inhaltlich als Wirkung von Bildungsaktivitäten beobachtet werden kann. In Anlehnung an Guskey (2002) zum Beispiel lassen sich folgende fünf Bereiche unterscheiden:

- die Reaktion der Zielgruppe(n) (wie beispielsweise die Zufriedenheit von Teilnehmenden mit der Maßnahme an sich, mit einzelnen Inputs oder mit verwendeten Lehr-Lernmaterialien; Bewertungen von im Rahmen der Maßnahme erstellten Produkten; themenbezogener Interessenszuwachs bei den Teilnehmenden nach einer Bildungsveranstaltung etc.)
- das Lernen der Zielgruppe(n) (selbstberichtete oder beobachtete Lernprozesse, die als solche beschrieben werden, ohne dass von diesen auf ‚Bewusstsein‘ geschlossen wird),

- die Anwendung des Gelernten (durch Beobachtung oder Testung unter bestimmten vorgegebenen Bedingungen oder in der alltäglichen Praxis),
- der Auf- und/oder Ausbau von Unterstützungsstrukturen sowie
- wahrnehmbare Veränderungen im (Projekt-)Umfeld (zum Beispiel Breitenwirksamkeit; gestiegene Nachfrage nach einem bestimmten Material/Angebot oder einer bestimmten Dienstleistung; Partizipation bisher unbeteiligter Personen an bestimmten Teilaspekten eines Projektes; organisationale Veränderungen bei den Bildungsakteuren selbst etc.).

Jeder dieser Bereiche kann in Wirkungsbeobachtungen im Globalen Lernen in den Blick genommen werden. Zu Rechenschafts- oder Entwicklungszwecken sollten aus diesen Bereichen allerdings diejenigen ausgewählt werden, die möglichst unmittelbar dem eigenen Einfluss unterliegen. Reaktionen der Zielgruppe beispielsweise lassen sich sowohl nach einem 45-Minuten-Input erfragen als auch nach einem vierwöchigen Bildungsprojekt. Gleiches gilt für eine Zunahme an Interesse oder Wissen. Kompetenzen wiederum lassen sich in einer kurzen Unterrichtseinheit nur sehr begrenzt so aufbauen, dass sie unmittelbar der jeweiligen Veranstaltung zugerechnet werden können, denn der Aufbau von Kompetenzen benötigt Zeit und entsprechende Lernformen, die bei kurzen pädagogischen Einsätzen meist nur schwer darzubieten sind. Gleiches gilt für den Aufbau von Haltungen (vgl. hierzu weiterführend Bergmüller & Paasch 2008). Was im Einzelnen als direkte bzw. indirekte Wirkungen beschrieben werden kann, ist immer abhängig von der jeweiligen Maßnahme und somit weitgehend fallspezifisch zu entscheiden. Deshalb haben die augenblicklich diskutierten Wirkungsketten (vgl. z.B. Massing et al. 2010, S. 54 ff.) auch immer nur Beispielcharakter.

In der schulischen Evaluationsforschung werden in Wirkungsbeobachtungen zumeist nicht nur die intendierten Wirkungen in den Blick genommen. Um zu einer möglichst korrekten Aussage über Ursache-Wirkungszusammenhänge zu kommen, werden in vielen Forschungsprojekten auch zudem Faktoren in den Blick genommen, die als (bekannte oder noch unbekannt) Kontextfaktoren wiederum den eigenen Einfluss beeinflussen können, so z.B. die Lernvoraussetzungen, Vorerfahrungen oder Intelligenz. Für viele Bildungsprojekte im Globalen Lernen ist dies angesichts der knappen zeitlichen, finanziellen und personellen Ressourcen kaum leistbar und im Sinne effektiven Ressourceneinsatzes auch nicht immer sinnvoll. Hier ist es aber deshalb sinnvoll, sehr genau zu benennen, was beobachtet wurde und die eigene Evaluation nicht überzustrapazieren.

#### **Anregungen für das Verständnis von Ursache-Wirkungsrelationen**

Wer über Ursache-Wirkungsrelationen sprechen möchte, sollte diese zu Beginn einer Evaluation dezidiert reflektieren und darstellen. Dabei ist alleine schon auf die verwendete Begrifflichkeit zu achten. Wer von einer „Kette“ spricht, unterstellt, dass es ein verbindendes Band gibt, auf die die einzelnen Beobachtungen sich wie die Glieder auf ein Kettenband aufreihen. Es sollte kritisch reflektiert werden, ob die in Wirkungsketten dargestellten Ursache-Wirkungszusammenhänge für die jeweils

intendierte Analyse von Wirkungen komplex genug sind oder ob nicht komplexere Wirkungsmodelle zu Grunde gelegt werden müssten – nicht unbedingt für die Planung, in jedem Fall aber, wie bereits dargestellt, für die Auswertung von Bildungsprozessen. Ein ausreichend differenziertes Wirkungsmodell würde es erleichtern, zu entscheiden, welche Daten insgesamt zur validen Einschätzung von Ursache-Wirkungszusammenhängen erhoben werden müssen bzw. welche Zusammenhänge (z.B. aufgrund fehlender Ressourcen oder fehlender Zugangsmöglichkeiten zum Feld) nicht überprüft werden können. Zusammenhänge nicht über Daten zu überprüfen heißt nicht, diese nicht im Blick zu haben; denn vielmehr lassen sich Zusammenhänge durch einen Rückbezug auf ein Modell dennoch – wenngleich dann als Vermutungen – mitreflektieren. Beim Umgang mit Wirkungsmodellen sollte grundsätzlich nicht übersehen werden, dass bei jeder Darstellung von Ursache-Wirkungszusammenhängen Komplexität reduziert wird und damit auch bestimmte Zusammenhänge bewusst oder unbewusst ausgeblendet werden. Die Rückführung von Veränderungen auf bestimmte Maßnahmen hat also aufgrund der nicht abbildbaren Vielschichtigkeit von Wirkungszusammenhängen immer einen gewissen Interpretationsstatus. Die getroffenen Zuordnungen könnten, von einem anderen Standpunkt aus gesehen oder in einen anderen oder differenzierten (Beobachtungs-)Kontext eingebettet, eventuell auch anders aussehen. Dieser Interpretationsstatus sollte bei Wirkungsbeobachtungen deutlich gemacht werden. Hilfreich – und auch als Standard von Evaluationen formuliert – ist es deshalb, die Trennung zwischen den Befunden als Ergebnisdarstellung und Interpretation der Ergebnisse auch bei der Beobachtung und Analyse von Wirkungen im Feld der entwicklungsbezogenen Bildungsarbeit zu berücksichtigen: In diesem Sinne müsste der Fokus der Wirkungsbeobachtung zunächst auf die rein deskriptive Darstellung von ‚Veränderungen‘ gerichtet werden, wobei mit ‚Veränderung‘ zunächst – objektiv – der Unterschied zwischen einem Ausgangszustand und dem Zustand zum Beobachtungszeitpunkt gemeint ist. Die (subjektive) Bewertung dieser Veränderung als ‚Wirkung‘ einer bestimmten Maßnahme müsste dann – deutlich von der Beschreibung der Veränderung getrennt – in einem zweiten Schritt erfolgen. Damit wird (vor allem auch Außenstehenden) klar, dass es sich bei dieser Bewertung um die Perspektive des- oder derjenigen handelt, der/die diese Bewertung vornimmt und diese Bewertung somit auch diskutiert werden kann.

#### **Anregungen für die Erhebung von Daten**

bleiben zum Schluss noch Anregungen für die Datenerhebung: Die Gefahr, sich mit zu umfangreichen Datenerhebungen zu überfordern, ist groß. Wichtig ist daher, genau zu überlegen, welche Datenerhebungsmethoden für die jeweilige Wirkungsbeobachtung sinnvoll und machbar sind. Prinzipiell steht den Akteuren zum einen die gesamte Bandbreite sozialwissenschaftlicher Datenerhebungsinstrumente zur Verfügung. Entsprechend der oben genannten Wirkungsbereiche könnten zum Beispiel

- für die Reaktion der Zielgruppe(n) schriftliche/mündliche Befragungen oder Kurzfeedbacks,
- für das Lernen der Zielgruppe(n) (selbst berichtet oder tatsächlich) Befragungen, Lerntests, Portfolios oder retrospektive Erzählungen,

- für die Anwendung des Gelernten (unter bestimmten vorgegebenen Bedingungen oder in der alltäglichen Praxis) (teilnehmende) Beobachtungen, Simulationsspiele, Videografien oder Tonbandaufzeichnungen sowie
- für wahrnehmbare Veränderungen im Umfeld Interviews, Fragebögen, statistische Daten oder die Auswertungen von Medienberichten eingesetzt werden.

Wichtig ist es, diese Datenerhebungsmethoden in ein angemessenes Gesamtkonzept zu bringen. Da es bei Wirkungsbeobachtungen um die Analyse von Veränderungen geht, kann zwischen folgenden Vorgehensweisen gewählt werden: 1) der Erhebung retrospektiver Einschätzungen, 2) einem Vorher-Nachher-Vergleich oder 3) einem Kontrollgruppenvergleich zwischen der jeweiligen Teilnehmergruppe und einer Gruppe, die in ihrer Zusammensetzung mit der Teilnehmergruppe vergleichbar ist, aber nicht am jeweiligen Bildungsprojekt teilgenommen hat.

Welche Datenerhebungsmethode und welches Gesamtkonzept die bzw. das richtige ist, ist abhängig vom jeweiligen Erkenntnisinteresse, den Möglichkeiten, die Methode überhaupt einzusetzen sowie dem eigenen Know-how, die mit der Methode erhobenen Daten auch auszuwerten und den zur Verfügung stehenden Ressourcen. Unter Rückgriff auf ein Wirkungsmodell kann gelten: Besser weniger Daten erheben, diese aber fundiert auswerten und die nicht erhobenen Zusammenhänge in der Datenauswertung offen thematisieren. In einem Bericht ließe sich dies beispielsweise folgendermaßen formulieren: ‚Eine Vorher-Nachher-Erhebung von Lernzuwachsen war aufgrund des Projektverlaufs nicht möglich, daher wurde auf retrospektive Einschätzungen der Teilnehmenden zurückgegriffen. Damit können zwar nur bedingt Aussagen über tatsächliche Lernzuwächse gemacht werden, in jedem Fall aber liefert die Wahrnehmung der Teilnehmenden Erkenntnisse darüber, inwiefern die vermittelten Inhalte für sie Bedeutung erhalten haben.‘

Natürlich können retrospektive Einschätzungen und Vorher-Nachher-Erhebungen oder Kontrollgruppenvergleiche auch kombiniert werden. Ein prominentes Beispiel für eine Kombination von retrospektiver Einschätzung und Vorher-Nachher-Erhebung bietet z.B. die Evaluation des BLK-Programms „21“ zu Motivation, Transfer und Gestaltungskompetenz (vgl. Rode 2005).

### Schlussbemerkung

Es ist davon auszugehen, dass das Erstarken von Wirkungsorientierung in der entwicklungsbezogenen Bildungsarbeit zu einem flächendeckenden Kulturwandel führen wird. Die eingangs geschilderten Herausforderungen sowie sicherlich noch zahlreiche weitere neuralgische Punkte, die bei Wirkungsbeobachtungen im Globalen Lernen auftreten, machen deutlich, dass in der Szene Globalen Lernens somit ein wachsender Professionalisierungsbedarf besteht, mit diesen Herausforderungen auch angemessen umgehen zu können. In diesem Zusammenhang wird es vor allem darum gehen, den Bereich der Selbstevaluation weiter zu stärken. Im Rückblick auf die letzten Jahre ist es bemerkenswert, wie sehr hier bei aller Ressourcenknappheit um Professionalität gerungen und welche Bedeutung dem Thema Wirkungsbeobachtung beigemessen wird. Diesem Rin-

gen gilt es, mit entsprechenden Fortbildungsmöglichkeiten zu begegnen. Hier wird es vor allem wichtig sein, Lösungsmöglichkeiten dafür aufzeigen, wie trotz knapper Ressourcen die Qualitätsansprüche an Wirkungsbeobachtungen erfüllt werden können.

### Anmerkung:

- 1 Dieser Beitrag stellt eine erweiterte Fassung des Beitrags „Bergmüller, C. (2012): Lehrkräfte und NROs als Qualitätsmanager des Globalen Lernens. Das Potenzial der Wirkungsbeobachtung im Globalen Lernen für die Zusammenarbeit von Schulen und NROs, in VENRO (Hg.): Jahrbuch Globales Lernen 2012. Wirkungsbeobachtung und Qualitätsentwicklung. Erfahrungen – Herausforderungen – Perspektiven. Bonn“ dar.

### Literatur:

- Bergmüller, C. (2012):** Schulen und NROs als Qualitätsmanager des Globalen Lernens. In: VENRO (Hg.): Jahrbuch Globales Lernen 2012. Wirkungsbeobachtung und Qualitätsentwicklung. Erfahrungen – Herausforderungen – Perspektiven. Bonn.
- Bergmüller, C./Paasch, D. (2008):** Evaluation von Lerneffekten – Potenziale und Grenzen. In: Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik, Jg. 31, H. 2, S. 15–19.
- Bergmüller, C./Scheunpflug, A. (im Erscheinen):** Die Evaluation entwicklungsbezogener Bildungsarbeit. In: Lang-Wojtasik, Gregor; Klemm, Ulrich (Hg.): Handlexikon Globales Lernen. Ulm.
- Bergmüller, C./Stadler-Altman, U. (2009):** Evaluationsforschung in der Schule. In: Blömeke, S./Bohl, Th./Haag, L./Lang-Wojtasik, G./Sacher, W. (Hg.): Handbuch Schule. Bad Heilbrunn, S. 153–156.
- Bundesministerium für wirtschaftliche Zusammenarbeit (BMZ) (2007):** BMZ Konzepte 159. Entwicklungspolitische Informations- und Bildungsarbeit. (Online im Internet: <http://www.bmz.de/de/publikationen/reihen/strategiepapiere/konzept159.pdf> [10.10.2011]).
- Bundesministerium für wirtschaftliche Zusammenarbeit BMZ (2008):** „Vorgaben nach den Allgemeinen und Besonderen Nebenbestimmungen für Zuwendungen des BMZ für die Förderung der Entwicklungspolitischen Bildung (ANBest-P/EB)“. Schreiben vom 19.09.2008.
- Deutsche Welthungerhilfe e.V. (2008):** Leitfaden Wirkungsorientierung in den Projekten und Programmen der Welthungerhilfe. Teil I: Hintergründe und Definitionen. Bonn.
- Engagement Global/Förderprogramm Entwicklungspolitische Bildung (FEB) (2011):** Verwendungsnachweis. (Online im Internet unter: <http://www.engagement-global.de/feb-abrechnung.html>, [15.01.2012]).
- Fend, H. (1998):** Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung. Weinheim: Juventa.
- Gohl, E. (2002):** prüfen und lernen. Praxisorientierte Handreichung zur Wirkungsbeobachtung und Evaluation. Bonn. VENRO.
- Guskey, T. R. (2002):** Does It Make a Difference? Evaluating Professional Development. In: Educational Leadership. Volume 59 Number 6, S. 45–51.
- Helmke, A. (2009):** Unterrichtsqualität und Lehrprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. Seelze-Velber.
- Joint Committee for Standards in Educational Evaluation (2011):** Classroom Assessment Standards: How to Improve Assessment Practices in K-12 Education. Draft #3. (Online im Internet: <http://www.jcsee.org/standards-development> [13.11.2011]).
- Joint Committee on Standards for Educational Evaluation, Sanders, J. R. (Hg.) (2006):** Handbuch der Evaluationsstandards: Die Standards des „Joint Committee on Standards for Educational Evaluation“. Wiesbaden.
- Jungk, S. (2010):** Mühen und Chancen der Evaluation – Zur Schwierigkeit der Messung von Wirkung in der Bildungsarbeit des Globalen Lernens. In: Massing, Armin; Rosen, Andreas; Struck, Gabi: Wirkt so. Handreichung zur Wirkungsorientierung in der Antragstellung in der entwicklungspolitischen Inlandsarbeit. Berlin, S. 18–23.
- Luhmann, N./Schorr, K.E. (1979):** Das Technologiedefizit der Erziehung und die Pädagogik. In: Zeitschrift für Pädagogik, 25, 3, S. 344–364.
- Massing, A./Rosen, A./Struck, G. (2010):** Wirkt so. Handreichung zur Wirkungsorientierung in der Antragstellung in der entwicklungspolitischen Inlandsarbeit. Berlin.
- Merkelbach, V. (2003):** Können unsere Schulen besser werden? Bildungsstandards und Tests im gegliederten Schulsystem. (Online im Internet: <http://user.uni-frankfurt.de/~merkelba/standards.htm> [10.10.2011]).

**OECD (2009):** Glossary of Key Terms in Evaluation and Results Based Management. (Online im Internet: <http://www.oecd.org/dataoecd/40/7/43184177.pdf> [15.3.2012]).

**Porter Michael (1980):** Competitive Strategy: Techniques for analyzing industries and competitors. New York.

**Prengel, A./Heinzel, F./Carle, U. (2004):** Methoden der Handlungs-, Praxis- und Evaluationsforschung. In: Helsper, W. & Böhme, J.: Handbuch der Schulforschung. Wiesbaden. S. 183–199.

**Rode, H. (2005):** Motivation, Transfer und Gestaltungscompetenz Ergebnisse der Abschlussequation des BLK-Programms „21“ 1999–2004. Berlin.

**Rosen, A. (2010):** Keine Wirkung, kein Problem? In der Wirkungsorientierung liegen Chancen für eine bessere entwicklungspolitische Inlandsarbeit. In: Massing, A.; Rosen, A.; Struck, G. (Hg.): Wirkt so. Handreichung zur Wirkungsorientierung in der Antragstellung in der entwicklungspolitischen Inlandsarbeit. Berlin. S. 10–13.

**Scherb, A. (2010):** Der Beutelsbacher Konsens. In: Lange, D./Reinhardt, V. (Hg.): Strategien der politischen Bildung. Handbuch für den sozialwissenschaftlichen Unterricht. Baltmannsweiler. Hohengehren (Basiswissen politische Bildung, Bd. 2), S. 31–39.

**Scheunpflug, A. (2004):** Das Technologiedefizit. Nachdenken über Unterricht aus systemtheoretischer Perspektive. In: Lenzen, Dieter (Hg.): Irritationen des Erziehungssystems. Pädagogische Resonanzen auf Niklas Luhmann. Frankfurt/Main. S. 65–87.

**Scheunpflug, A./Bergmüller, C./Schröck, N. (2010):** Evaluation entwicklungbezogener Bildungsarbeit. Eine Handreichung. Münster u.a.

**Seitz, K. (2010):** Aus Erfahrung lernen – Wirkungsanalyse der entwicklungspolitischen Inlandsarbeit bei „Brot für die Welt“. In: Massing, Armin; Rosen, Andreas; Struck, Gabi: Wirkt so. Handreichung zur Wirkungsorientierung in der Antragstellung in der entwicklungspolitischen Inlandsarbeit. Berlin. S. 14–17.

**Stiftung Nord-Süd-Brücken (2011):** Soll-Ist-Vergleich Personalkostenförderung „Multiplikator/innenausbildung“. (Online im Internet: [http://www.nord-sued-bruecken.de/assets/files/muster\\_fuer\\_Soll-Ist-Vergleich.pdf](http://www.nord-sued-bruecken.de/assets/files/muster_fuer_Soll-Ist-Vergleich.pdf). [10.3.2012]).

**VENRO (2010):** Qualität statt Beweis. VENRO Positionspapier 2/2010 zur Wirkungsbeobachtung. Bonn.

### Dr. Claudia Bergmüller

ist wissenschaftliche Mitarbeiterin am Lehrstuhl für Allgemeine Erziehungswissenschaft I der Friedrich-Alexander-Universität Erlangen-Nürnberg; Arbeitsschwerpunkte u.a.: Qualitativ-rekonstruktive Bildungs- und Evaluationsforschung, wissenschaftliche Begleitung/Evaluation von Bildungsprojekten im Inland, Evaluation von Bildungsprojekten in der Entwicklungszusammenarbeit; seit 2003 Referentin für Evaluation entwicklungbezogener Bildungsarbeit.