

Frühjahrstagung zum Thema Relevanz des AK-Entwicklungspolitik und Humanitäre Hilfe in Bad Homburg

Relevanz – Tun wir die richtigen Dinge?

„Almost every problem confronting our society is a result of the fact that our public policy makers are doing the wrong things and are trying to do them righter.“ Russell Ackoff (2004: 2)

Maximilian Schmid,¹ Amy O'Brien,² Thorsten Bär,² Susanne von Jan¹

Tun wir die richtigen Dinge? Zielen die Maßnahmen auf die Lösung der Kernprobleme der Zielbevölkerung? Inwieweit stimmt die Zielsetzung des Projekts mit der des Partnerlandes und den Strategien und Richtlinien des Mittelgebers überein? So lauten grundlegende Fragen des BMZ für eine Bewertung des OECD/DAC-Kriteriums Relevanz. Die genaue Umsetzung des Relevanzkriteriums in der Entwicklungszusammenarbeit wie auch in der Humanitären Hilfe (Relevanz/Angemessenheit) wirft jedoch Fragen auf: Wie kann das Kriterium systematisch untersucht werden und welche Fragestellungen und Potenziale beinhaltet das Kriterium im Detail? Wie können mögliche projektrelevante Veränderungen analysiert werden und wie erfolgt die Überprüfung, ob die Aktivitäten und Outputs konsistent zu den Zielen des Projektes sind, ob die Wirkungsannahmen also ausreichend plausibel sind? Und um wessen Bedarfe geht es, wenn man die Relevanz eines Projekts analysieren möchte – die Relevanz aus Sicht der Zielgruppe, aus Sicht des Partnerlandes, aus Sicht des Gebers?

All diesen Fragen sollte nun auf der Frühjahrstagung des Arbeitskreises Entwicklungspolitik und Humanitäre Hilfe (DeGEval) im April 2018 im hessischen Bad Homburg nachgegangen werden. Etwa 90 Teilnehmende folgten der Einladung zur Tagung in das Schloss Bad Homburg, wo sie die GOPA Consulting Group als Gastgeberin empfing. In ihrer Begrüßungsrede griffen GOPA-Geschäftsführer Berthold Averweg und die Leiterin des Bereiches für Monitoring & Evaluation/Quali-

1 Hanns-Seidel-Stiftung e.V., München

2 World Vision Deutschland e.V., Friedrichsdorf

ty, Heike Happerschoss, die Frage auf, wie relevant die Relevanz nun sei und freuten sich auf eine ausführliche Behandlung des Kriteriums. Die Tagung verfolgte das Ziel, ein gemeinsames Verständnis bei den Teilnehmenden darüber zu erlangen, welche Inhalte und Leitfragen in das Relevanz-Kapitel gehören, wie das Kriterium zu anderen DAC-Kriterien abgegrenzt werden kann und welche Quellen zur Beurteilung herangezogen werden können. Ebenso war eine Darstellung und Diskussion der angelegten Bewertungskriterien bei der Veranstaltung vorgesehen.

Die DAC-Kriterien im Zeitalter der SDGs

Die Agenda der Tagung begann mit einem Vortrag zu den fünf DAC-Kriterien im Zeitalter der Agenda 2030 und der Ziele für Nachhaltige Entwicklung (SDGs). Bevor das DAC-Kriterium der Relevanz spezifisch betrachtet wurde, gaben Jörg Faust und Ida Verspohl vom Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit (DEVal) zuerst einen Überblick über aktuelle Diskussionen im Zusammenhang mit allen fünf DAC-Kriterien. Mit ihrem Beitrag verschafften sie den Teilnehmenden einen Überblick zu den Evaluierungskriterien und verwiesen sowohl auf deren Stärken als auch auf deren Herausforderungen. Das kohärente Set an Kriterien bietet eine Referenz- und Orientierungsfunktion und trägt mit seiner Lern- und Vergleichsfunktion zu einer Harmonisierung der EZ-Evaluierung im internationalen Kontext bei. Die Frage, wie man Relevanz prüft, erweist sich aufgrund der Vielzahl an verschiedenen Referenzsystemen als schwierig. Die Referent(inn)en des DEVals stellten die Prüffragen des BMZ vor und verwiesen auf Überlappungen mit den sogenannten 3K (Kohärenz, Komplementarität und Koordination) aus den Evaluierungskriterien der Humanitären Hilfe, die das BMZ auch für die Entwicklungszusammenarbeit heranzieht. Darüber hinaus bezogen Faust und Verspohl noch Stellung zum Nachhaltigkeitskriterium, was in seiner derzeitigen Konzeption auf die Messung der Dauerhaftigkeit der Wirkung reduziert werden könnte. Zusätzlich könnte ein eigenes Kriterium der „nachhaltigen Entwicklung“ eingefügt werden, das Bezug auf die Dimensionen der Nachhaltigkeit und die Kernprinzipien der SDGs nehmen sollte. Nachhaltigkeit sollte man, so die Referent(inn)en, als ein mehrdimensionales Thema betrachten.

Den Anschluss machte der Diskutant Bernward Causemann (*Causemann Consulting*) mit der Präsentation seiner Thesen, welche eine kritische Betrachtung der derzeitigen Verfahrensweisen mit den DAC-Kriterien beinhalten. Causemann kritisiert, dass die Kriterien tendenziell zu einer linearen Denkweise verleiten und eine tiefgehende sowie detaillierte Auseinandersetzung mit einzelnen wichtigen Fragen aufgrund der Abarbeitung aller Kriterien nur noch selten stattfindet. Zudem führt der Diskutant aus, dass die Beurteilung der Qualität der Umsetzung stark vernachlässigt wird. Es sei entwicklungspolitisch von höchster Bedeutung, dass aktuelle Standards und Expertenwissen angemessen genutzt und bestimmte Werte wie Partizipation oder demokratische Führung beachtet werden. Daraus erschließt sich Causemanns Forderung nach einem zusätzlichen Qualitätskriterium. Weiterhin empfiehlt Causemann die Kriterien Effektivität und Impact zusammenzufassen und auf das

Effizienzkriterium in small-e-Evaluierungen (Evaluierungen einzelner Projekte) zu verzichten.

Am Ende der beiden gehaltenen Vorträge wurde mit dem Plenum darüber diskutiert, ob die Forderung nach einer umfassenden Infragestellung der DAC-Kriterien oder nach einer grundsätzlichen Beibehaltung lediglich mit inhaltlichen Verfeinerungen der richtige Weg sei.

Das Relevanzkriterium in der EZ und HuHi

Der Nachmittag des ersten Tagungstages begann mit einem Austausch der Teilnehmenden zum Tagungsthema. Aufgeteilt in 12 Gruppen diskutierten die Teilnehmenden in Stationengesprächen vorbereitete Leitfragen, wie z.B.:

- Welche Leitfragen sollen im Relevanz-Kapitel beantwortet werden, um welche Inhalte geht es (auch in Abgrenzung zu den anderen DAC-Kriterien)?
- Welche Akteure sollten für die Erörterung der Relevanz eines Projekts in den Blick genommen werden (Stichworte: leave no one behind, Mehrebenenansatz)?
- Wie geht ein Gutachter bzw. eine Gutachterin am besten mit Widersprüchen um, wenn bestimmte Gruppen ein Projekt beispielsweise als sehr relevant erachten, andere als gar nicht relevant?
- Wie gehen Gutachter(innen) mit eigenen Grundpositionen um? Inwieweit gehen eigene Überzeugungen des Gutachters bzw. der Gutachterin in die Bewertung der Relevanz ein?
- Wie geht man an die Frage heran, ob andere Maßnahmen oder Aktivitäten besser für die Lösung der relevanten Probleme des Landes bzw. der Zielgruppen geeignet gewesen wären?
- Wie kann untersucht werden, ob die Aktivitäten und Maßnahmen für die lokalen Bedarfe angemessen waren (Stichwort: Appropriateness)? Wie gelingt die Abgrenzung zu Effektivität?

Im Anschluss führten Elias Sagmeister (Ground Truth Solutions) und Neil Dillon (ALNAP) mit ihren Vorträgen „*Evaluation von Relevanz in der EZ: Praxis und Reformbedarf*“ und „*Relevance and Appropriateness in Evaluation of Humanitarian Action*“ in die Betrachtung des Relevanzkriteriums im Feld der Entwicklungszusammenarbeit und der Humanitären Hilfe ein.

Elias Sagmeister, der zum Thema Relevanz promovierte, stellte in seinem Vortrag die Frage: Erfahren wir durch die durchgeführten Evaluationen, ob wir die richtigen Dinge tun? Einführend legte der Referent dar, dass Relevanz ein multifaktorielles Konstrukt ist: die Ziele müssen mit einer Vielzahl an Aspekten übereinstimmen (z.B. Strategien der Geber, Bedürfnisse der Zielgruppe etc.). Darüber hinaus müsse aber auch die Angemessenheit der untersuchten Maßnahmen betrachtet werden (z.B. in Bezug auf das Timing). Im Rahmen seiner Promotion führte Sagmeister u.a. eine Metaevaluation durch, in der er eine Zufallsstichprobe von knapp 80 Evaluierungsberichten von Entwicklungsorganisationen untersuchte. Dabei analysierte er, welche Aspekte in Evaluationen zur Beurteilung der Relevanz und der Angemessen-

heit herangezogen werden und welche Qualität die gängige Evaluationspraxis bei der Bewertung der Relevanz aufweist. Seine Ergebnisse zeigen, dass ein Großteil der berücksichtigten Evaluationen Relevanz meist unvollständig und oberflächlich betrachtete. Zudem falle die Relevanzbewertung in den meisten Fällen positiv aus. Reformbedarf sieht Sagmeister primär darin, dass die Perspektiven der Betroffenen priorisiert werden sollten und lokale Expert(inn)en stärker einbezogen werden müssten. Ebenso ist eine adäquate Verknüpfung der DAC-Kriterien erstrebenswert. Darüber hinaus sollten Standards für einen glaubwürdigen Analyseprozess festgelegt werden. Dies beinhaltet im Allgemeinen, dass das Selbstverständnis von Evaluationen überdacht werden müsse.

Neil Dillon vom ALNAP-Netzwerk ging folgend auf die Unterschiede der DAC-Kriterien zwischen der Humanitären Hilfe und der Entwicklungszusammenarbeit ein. In der Humanitären Hilfe wurde das Set um die 3C (Coherence, Connectedness, Coverage) erweitert. Dagegen ist die Analyse der Nachhaltigkeit prinzipiell nur im EZ-Bereich vorgesehen. Das Kriterium der Relevanz wiederum wird in der Humanitären Hilfe mit der Angemessenheit (Appropriateness) zusammengefasst. Zu Letzterem führt der Referent aus, dass die Ziele in einem Projekt zwar relevant sein können, dies aber nicht garantiert, dass die Aktivitäten auch angemessen sind. Bei der Durchführung einer Relevanzevaluierung seien nach Dillon vier elementare Fragen zu stellen:

- Erfüllen die Projektaktivitäten die prioritären Bedürfnisse der betroffenen Menschen?
- Erfüllt man die Bedürfnisse vulnerabler Bevölkerungsgruppen (Stichwort: leave no one behind)?
- Passen sich die Projektaktivitäten an die sich verändernden Bedürfnisse der Menschen an?
- Sind die Maßnahmen im Krisenkontext angemessen?

Bei der Analyse von 120 Evaluationsberichten aus den Jahren 2015 bis 2018 hinsichtlich dieser vier Fragen fiel ebenfalls auf, dass die Relevanz mehrheitlich mit gut bewertet wird. Allerdings wird die Relevanz laut Dillon vornehmlich anhand von Expertenmeinungen bewertet. Nur bei einer marginalen Zahl an Evaluierungen wird die Zielgruppe selbst befragt. Dillon resümiert, dass es drei Faktoren gibt, die eine zentrale Rolle für die Beurteilung von Relevanz spielen: Sind Bedarfsanalysen und Strategien vorhanden und wurde die Nutzung von Cash geprüft? Cash könnte die Relevanz erhöhen, da die Entscheidung, was am wichtigsten ist, so zum Rezipienten/zur Rezipientin übergeht. Dies wird von den Betroffenen aber nicht immer präferiert.

Sagmeister und Dillon standen zum Abschluss des ersten Tages in einer Fishbowl-Diskussion den Teilnehmenden der Konferenz Rede und Antwort. Hierbei wurde aus jeder Gruppe der zu Beginn des Tages stattgefundenen Stationengespräche, eine Frage an die Referenten gestellt. Zum Abschluss des ersten Tages betonten die Referenten, dass das Relevanzkriterium die Basis jeder Evaluation sei. Wenn die Relevanz eines Projektes nicht gegeben ist, dann sei eine gute Bewertung der anderen DAC-Kriterien irrelevant.

Wie steht es um die Relevanz der Relevanz?

Der zweite Tag der Konferenz begann mit einem Vortrag von Stefan Silvestrini der CEval GmbH mit dem Titel „*Die Relevanz von Relevanz – Beobachtungen aus sieben Meta-Studien zu Stellenwert, Analyse und Bewertung des OECD-DAC Kriteriums ‚Relevanz‘ in Evaluationen*“. In den sieben Metastudien wurden insgesamt ca. 250 Evaluationen von staatlichen und nicht staatlichen Organisationen analysiert. Wie bereits am Vortrag Dillon, konstatierte auch Silvestrini, dass die Relevanzbewertung in der Regel auf Experteneinschätzungen basiere. Spezifische Erhebungsinstrumente ließen sich für die Bewertung der Relevanz nicht identifizieren. Bei den Verfahrensweisen falle zudem auf, dass bei den Auswertungen fast ausschließlich qualitative Verfahren zum Einsatz kommen und eine Daten- und Quellentriangulation nur selten stattfindet. Der Referent bestärkt zudem den Kritikpunkt aus einigen der vorangegangenen Präsentationen, dass die Bedarfe der Endbegünstigten bei den Erhebungen zu wenig beachtet werden. Außerdem sei es auffallend, dass Relevanz das am besten bewertete DAC-Kriterium ist.

Anschließend standen zwei Werkstattgespräche zur Auswahl. Im ersten Werkstattgespräch wurde durch Harriet Lange von Plan International e.V. zuerst eine Evaluation mit dem Fokus auf eine spezifische Erhebungsmethode präsentiert. Die Bewertung der Relevanz eines Projektes erfolgte mit der People First Impact Method (P-FIM), welche darauf abzielt, Herausforderungen und vorhandene Lösungsansätze der Zielgruppen durch eine zielfreie Diskussion zu identifizieren. Lange führt aus, dass dem Projekt in einer vorhergehenden Evaluation eine hohe Relevanz attestiert wurde. Die zu einem späteren Zeitpunkt mit P-FIM erhobenen Daten präzisierten und ergänzten die Empfehlungen der früheren Evaluation. Demnach sei der Projektansatz grundsätzlich relevant, aber es wurden auch Empfehlungen zur thematischen Anpassung des Projektes gemacht, die seitens der Referentin als sehr nützlich bewertet wurden. Die P-FIM-Methode sei für die Bewertung der Relevanz eines Projektes besonders geeignet.

Das zweite Fallbeispiel zur Relevanzbewertung stellte eine gemeinschaftliche Evaluation der Maßnahmen verschiedener Mitgliedsorganisationen von Aktion Deutschland Hilft (ADH) in Haiti in Folge des Hurrikans Mathew vor. Die Evaluation sollte unter anderem die Relevanz und Angemessenheit der nach dem ersten Hurrikan 2010 eingeleiteten Katastrophenvorsorgemaßnahmen bewerten. Die Ergebnisse waren vorwiegend positiv. Der Referent Markus Moke (ADH) betonte die methodischen Herausforderungen von Evaluation im Kontext der Humanitären Hilfe. Teilweise waren ehemalige Projektmitarbeiter(innen) in den Organisationen nicht mehr verfügbar, was es erschwerte, ehemalige Zielgruppen ausfindig zu machen und Interviews zu organisieren. Auch erlaubte die Sicherheitslage und die mangelnde Infrastruktur nicht, alle Projektregionen zu besuchen. Dieses Beispiel zeigte anschaulich mögliche methodische Einschränkungen einer Evaluation im Kontext der Humanitären Hilfe, die eine Bewertung des Relevanzkriteriums erschweren.

In dem zweiten Werkstattgespräch präsentierten Ulrike Haffner und Benjamin Bräuer der Stabsstelle Evaluierung der Gesellschaft für Internationale Zusammenarbeit (GIZ) das neue Konzept der zentralen Projektevaluierungen (ZPE) der Orga-

nisation. Im Rahmen der Reform wurden unter anderem die Dimensionen zur Bewertung von Relevanz in Evaluationen angepasst. Hierbei sollen seit dem 1. Juni 2017 folgende Dimensionen bewertet und mit einem Punktsystem verknüpft werden: (1) die Relevanz bzgl. nationalen, internationalen und sektoralen Strategien, (2) die Relevanz bzgl. Bedarfen und Kernproblemen der Zielgruppe, (3) die Angemessenheit/Qualität der Projektkonzeption sowie (4) inwieweit ein Projektdesign aufgrund eventueller Veränderungen des Projektkontextes angepasst wurde. Als einen Diskussionspunkt, auf Basis der bisherigen Erfahrungen, warfen die Referenten auf, ob die Bewertung der Kohärenz der Projektkonzeption eher unter dem DAC-Kriterium Relevanz oder unter Effektivität/Impact betrachtet werden sollte.

Schlussrunde mit neuen und alten Erkenntnissen

Die Frühjahrstagung in Bad Homburg endete mit einer Podiumsdiskussion in der die Erkenntnisse der Konferenz zusammengefasst und diskutiert wurden. Die Podiumsgäste waren Heike Happerschoss (GOPA) für die Seite der Auftragnehmer, Anja Scherkus (Rosa-Luxemburg-Stiftung) für die Seite der Auftraggeber und Elias Sagmeister (Ground Truth Solutions).

Grundsätzlich wurde die Relevanz in der Diskussion einheitlich als notwendige aber nicht hinreichende Voraussetzung für ein gelungenes Projekt eingestuft. In der Diskussion wurde weiterhin festgehalten, dass sowohl in der Entwicklungszusammenarbeit als auch in der humanitären Hilfe sowohl die Übereinstimmungen der Ziele eines Projektes mit anderen Zielsystemen als auch die Angemessenheit eines Projektes behandelt werden sollte. Dabei sei es wichtig, zumindest die Bedürfnisse der Zielgruppe durch Befragungen direkt zu erfassen. Darüber hinaus sei es natürlich sinnvoll, auch die Einschätzungen aller anderen relevanten Stakeholder-Gruppen zu erheben. Es gelte dabei: Lieber weniger Personen pro Gruppe, als weniger Gruppen an sich zu befragen. Wie viele Personen befragt werden, hänge selbstverständlich auch vom Erkenntnisinteresse der Auftraggeber und der verfügbaren Ressourcen ab. Einig waren sich die Podiumsgäste, dass es besser sei, Erwartungen zu managen und Terms of References (ToR) zu reduzieren, als im Evaluierungsbericht oberflächliche Statements zu erhalten. Bei der Beurteilung der Angemessenheit sollte zumindest die Plausibilität der Wirkungslogik betrachtet werden, weitere Fragestellungen hängen ebenfalls vom Erkenntnisinteresse der Auftraggeber ab. Im Rahmen der Tagung wurde des Öfteren die unklare Trennschärfe zwischen den Kriterien, vor allen Dingen zwischen Relevanz und Effektivität, thematisiert. Hierzu wurde in der Schlussdiskussion festgehalten, dass es Querverbindungen gäbe, man jedoch bedenken müsse: Ziele können plausibel sein aber dennoch nicht erreicht werden, ebenso können Ziele erreicht werden, allerdings nicht relevant sein. Hinsichtlich der Beurteilung von Relevanz sei es wichtig, dass nicht die persönlichen Werte des Gutachters/der Gutachterin im Vordergrund stehen, sondern nachvollziehbare Bewertungskriterien erarbeitet und dargestellt werden. Sollten sich bei der Beurteilung der Relevanz durch die Stakeholder Widersprüche ergeben, so müssen diese Widersprüche offengelegt werden.

Schließlich dürfte den Teilnehmenden das gegen Ende geäußerte Plädoyer in Erinnerung bleiben: Man solle den Mut haben, Relevanz kritisch zu betrachten, denn es ist besser ein gutes Projekt schlecht zu machen, als ein schlechtes Projekt gut.

Literatur

Ackoff, Russell L. (2004): Transforming the Systems Movement. In: *The Systems Thinker*, 15 (8), S. 2-5.