

Bernward Causemann

Die DAC-Kriterien im Zeitalter der SDGs

Discussant-Beitrag: 8 Thesen

DeGEval, Arbeitskreis Entwicklungspolitik, Frühjahrstagung, 11.4.2018

1. Die DAC-Kriterien haben sowohl genutzt als auch geschadet. Verglichen mit Evaluierungen, die vor 15 Jahren gemacht wurden, als die DAC-Kriterien noch nicht allgemein verbreitet waren, gibt es heute weniger sehr schlechte, aber auch weniger sehr gute Evaluierungen. Dass es weniger sehr schlechte gibt, hat auch damit zu tun, dass es solche Kriterien gibt. Dass es weniger sehr gute gibt, hat damit zu tun, welche Kriterien es sind. Die DAC Kriterien müssen daher nicht nur immanent verbessert, sondern insgesamt in Frage gestellt werden.
2. Wir müssen unterscheiden zwischen „Evaluation“ (small-e, d.h. der Bewertung einzelner Projekte) und „Evaluationsforschung“ („Big-E“ – Forschung mit den Methoden der Evaluation) (nach Wollmann 2017, Whitty 2015). Ich orientiere mich beim Punkt 3 an der Praxis des small-e.
3. Die Kriterien werden im small-e oft zu brav umgesetzt: Es wird häufig nicht mehr gefragt, was interessiert und wichtig ist, sondern es wird überlegt, was gefragt werden müsste. So werden zu viele Fragen aufgelistet, und irrelevante Fragen gestellt. Das gilt für alle Kriterien, aber besonders für Effizienz, sowie in der Verwirrung von Effektivität und Impact – die 5-stufige Wirkungskette ist unangemessen für komplexe Vorhaben, die DAC-Kriterien verleiten zu einer linearen Denkweise (nicht zwingend, aber tendenziell). Es wird auch zu selten auf ein Kriterium oder auf eine Unterfrage in Kriterien verzichtet. Auch Gutachter*innen passen sich häufig zu stark an diese Struktur an. Die Kriterien beeinträchtigen die geistige Unabhängigkeit, was sich dann negativ auf die Ergebnisse auswirkt. Es gelingt häufig nicht genug, sich voll auf die Materie einzulassen, und dies in den Fragen und der Berichtsstruktur zum Ausdruck zu bringen. Big-E scheint von diesem Problem nach meiner Beobachtung nicht berührt zu sein. Dort scheint es zu gelingen, im Rahmen der Kriterien wesentliche interessante Fragestellungen zu entwickeln und zu bearbeiten.
4. Die Kriterien führen dazu, dass Zentrales vernachlässigt wird. Die Qualität der Umsetzung ist mindestens so wichtig wie Wirkung – in ihr zeigt sich das, was langfristig von höchster Bedeutung (und wirksam) ist, aber unter dem Kriterium Wirkung nicht erfasst werden kann. Dabei geht es um zweierlei: Zum einen die Frage, ob die Qualitätsansprüche des jeweiligen Fachgebiets, die Standards, das Expertenwissen angemessen genutzt und in den jeweiligen Kontext eingepasst wurden, d.h. ob Wasserprojekte, Mikrofinanzprogramme oder Maßnahmen der Politikbeeinflussung etc. Prinzipien von good practice im jeweiligen Sektor entsprechen. Ob die Zugrunde liegende Logik dieser Interventionen auch verstanden und aufgenommen wird, ist eine Bedingung dafür, dass langfristig eine vernünftige Inkulturation gelingt. Alle Entwicklungsmaßnahmen sind verstörende Interventionen, die aus einer fremden Kultur kommen. Es kommt darauf an, dass diese Störungen produktiv aufgenommen werden, und die kulturell sensible Beachtung von Expertenwissen ist eine Voraussetzung dafür, die sich weit über den Zeitraum des Projektes hinaus positiv oder, wenn die Inkulturation nicht gelingt, negativ auswirken kann. Zum anderen geht es um die Frage, ob bestimmte Werte, wie Partizipation, Öffnung für Lernen, demokratische Führung, etc. beachtet werden. Welche Werte zu beachten seien, ist vor der Evaluierung zu bestimmen. Die Capacity WORKS Erfolgsfaktoren erfassen das nur zum Teil. Auch die Beachtung sol-

cher Prinzipien ist langfristig wirksam. Solche Fragen können auch derzeit unter den DAC-Kriterien betrachtet werden. Es ist aber der Entscheidung der Beauftragenden und der Evaluierenden überlassen, an welchen Stellen und in welcher Ausführlichkeit diese Fragen der Qualität analysiert werden. Es hat praktische Folgen. So monieren Evaluierungen zivilgesellschaftlicher Projekte, die sich direkt an die Bevölkerung wenden, immer wieder die Qualität der M&E-Systeme. Dass nur ein kleiner Teil der Mitarbeiter*innen dieser Organisationen eine ordentliche Ausbildung in partizipativer Methodik hat, wird fast nie thematisiert. Dabei ist das zweite Problem unter Umständen folgenreicher als das erste. Das Fehlen eines Kriteriums Qualität ist ein Problem für die small-e wie für Big-E und für die einzelne Evaluierung ebenso wie für den Versuch, aus der Gesamtheit der Evaluierungen zu lernen. Querschnitts-Studien zu diesen Themen werden dadurch massiv behindert. Das Thema Qualität verdient mehr Fokus. Dafür sollte ein zusätzliches Kriterium Qualität eingeführt und inhaltlich ausgestaltet werden.

5. Effizienz sollte nur dann abgefragt werden, wenn das Thema relevant ist und sinnvolle Erkenntnisse zu erwarten sind. Ansonsten sind für Effizienz eher spezifische Big-E Studien sinnvoll, weil sie die Möglichkeit herstellen können, Vergleiche verschiedener Maßnahmen auf spezifische Interventionen anzustellen. Das ist tendenziell für Maßnahmen mit linearer „einfacher“ Wirkungslogik sinnvoller. Aus ihnen kann dann allgemein gelernt und dieses Wissen auf small-e Evaluierungen angewandt werden. Aus diesem Grund sollte Effizienz als eigenes Kriterium gestrichen und das Thema unter Qualität gefasst werden.

6. Zudem könnten die beiden Kriterien Effektivität und Impact, deren Abgrenzung und Bedeutung nach Wissenschaftstradition beträchtlich changiert, zusammengefasst werden, um dann Fragen spezifischer ausformulieren können. Dann blieben vier Kriterien über: Relevanz, Qualität, Effektivität/Impact sowie Nachhaltigkeit. Der Vorschlag aus dem DEval, das Nachhaltigkeits-Kriterium aufzuspalten in Dauerhaftigkeit und SDGs ist dabei nicht berücksichtigt.

7. Bisher kann man Evaluationen dann am besten qualitativ gestalten, wenn man mit den DAC-Kriterien subversiv umgeht, d.h. sie partiell unterläuft und z.B. Themen hineinschmuggelt, für die sie nur am Rande gedacht waren. Aber der Sinn der Vergleichbarkeit, für die die Kriterien geschaffen wurden, geht dabei verloren.

8. Von einer standardisierten Ausrichtung der Evaluation an den Zielen der SDGs ist abzuraten, weil sie aller Wahrscheinlichkeit zu Fehlsteuerungen führen wird. Da gilt Campbell's Law: "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor." (Campbell 1976) Das war schon bei den MDGs zu beobachten. Das Ziel, die Primarbildung auszuweiten, wurde an der Zahl der Primarschüler*innen gemessen. Als Konsequenz erhöhten manche Länder die Klassengröße. Der Bildungsprozess wurde dadurch beschädigt. Auch wenn das später z.T. korrigiert wurde – der Schaden hat ganze Schülergenerationen getroffen. Ähnlich beim MDG 1, das auf Einkommen setzte. Armut misst sich aber mehr in der Gefährdung durch Schocks (weswegen sie eher Risiko-advers sind; Berner u.a. 2011), wogegen Sparen hilft. Das MDG verstärkte die Strategie vieler Geber, vor allem auf Mikrokredite zur Armutsbekämpfung zu setzen. Eine Betonung der Reduktion von Vulnerabilität hätte stattdessen die Strategie in Richtung Sparen und Aufbau von Sozialkapital lenken können – vermutlich ein effektiverer Beitrag zur Reduktion extremer Armut. Ähnliche Schwachstellen haben die SDGs heute schon, und weitere werden im Laufe der Jahre folgen. Für eine zu enge Orientierung an ihnen wären deswegen ähnliche Fehlsteuerungen zu erwarten. Evaluierungen sollten möglichst frei sein, solches Wirken von Campbells Gesetz zu identifizieren. Eine Orientierung der Evaluation an den Prin-

zipien der SDGs lässt solche Nebenwirkungen weniger erwarten. Sie könnte jedoch auch im Kriterium Qualität unter den Werte-basierten Fragen geschehen.

Literatur:

- Berner, Erhard, Georgina Gomez, Peter Knorrinda (2011): Ein bisschen weniger arm. Die Logik der Subsistenzunternehmer, *Peripherie* 121, S. 7-26
- Campbell, Donald T (1976): Assessing the Impact of Planned Social Change, Paper 8, Occasional Paper Series, The Public Affairs Centre, Dartmouth College
- Eyben, Rosalind et al (ed) (2015): *The Politics of Evidence and Results in International Development. Playing the game to change the rules?* Rugby: Practical Action Publishing, developmentbookshop.com/the-politics-of-evidence-and-results-in-international-development
- Whitty, Brenden (2015): Mapping the results landscape: insights from a crowdsourcing survey, pp39-58, in: Eyben et al. 2015
- Wollmann, Hellmut (2017): Entwicklungslinien von Evaluation und Evaluationsforschung in Deutschland. Zwischen Vorgeschichte, Aufbruch und Konsolidierung, *ZfEv* 2/2017, S. 33-53