DEGEVAL … INFO

Zeitschrift für Evaluation | 24. Jahrgang | 2025
Heft 2 | S. 386–394 | https://doi.org/10.31244/zfe.2025.02.17 |
© 2025 Waxmann

Nachhaltigkeit evaluieren: zwischen Anspruch und Realität

Herausforderungen und Chancen für die Bewertung von Dauerhaftigkeit und Anschlussfähigkeit in Entwicklungszusammenarbeit (EZ) und humanitärer Hilfe – Bericht zur Frühjahrstagung des AK EPol und HuHi

Julian Friedrich[1]

Als eines der sechs Evaluierungskriterien der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) ist Nachhaltigkeit fest in Evaluierungen verankert. Gleichzeitig existieren bis heute unterschiedliche Interpretationen von Nachhaltigkeit und dadurch verschiedene Umsetzungen in der Evaluierung des Kriteriums. Deshalb standen auf der diesjährigen Frühjahrstagung des Arbeitskreises Entwicklungspolitik und Humanitäre Hilfe (AK Epol-HuHi) der Gesellschaft für Evaluation (DeGEval), die gemeinsam mit der GFA Consulting Group in Hamburg am 15. und 16. Mai 2025 ausgerichtet wurde, verschiedene Perspektiven und Auslegungen der Evaluierung von Nachhaltigkeit im Fokus der Debatte. Freie Evaluator:innen, Vertreter:innen staatlicher und nicht-staatlicher Organisationen sowie Mitarbeiter:innen aus der Privatwirtschaft diskutierten zwei Tage lang über die Evaluierung von Nachhaltigkeit bei Maßnahmen der EZ und humanitären Hilfe. In einer abschließenden Diskussionsrunde erfolgte eine Reflexion zu den Mittelkürzungen und den damit verbundenen Implikationen für Nachhaltigkeit von Entwicklungsmaßnahmen sowie deren Evaluierung.

Den Auftakt machte Claudio Alberti vom Evaluierungsnetzwerk des Entwicklungsausschusses der OECD (DAC EvalNet). In seinem Vortrag zeichnete er zunächst die Genese des Kriteriums innerhalb der OECD nach und stellte anschließend die ab 2019 erarbeitete Neudefinition vor: „Nachhaltigkeit bezeichnet das Ausmaß, in dem die Nettowirkungen der Maßnahme anhalten oder wahrscheinlich anhalten werden" (OECD, 2019, S. 12). Während die Initialdefinition von 1991 auf die Dauerhaftigkeit externer Entwicklungsmaßnahmen fokussierte, trägt die Neudefinition des Kriteriums der Multidimensionalität von Nachhaltigkeit stärker Rechnung und weist der Kontextualisierung eine größere Bedeutung zu. So erfordere die Bewertung von Nachhaltigkeit nun eine umfassende Analyse der finanziellen, ökonomischen, ökologischen, so-

1 DEval – Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit

zialen und institutionellen Kapazitäten der Systeme, die nötig wären, um die erzielten Nettowirkungen aufrechtzuerhalten (OECD, 2021, S. 71). Anschließend warf der Vortrag einen Blick auf die Anwendung des Kriteriums anhand von Evaluierungen der britischen Independent Commission for Aid Impact (ICAI, 2018), des Deutschen Evaluierungsinstituts der Entwicklungszusammenarbeit (DEval) (Orth et al., 2018) und der Japan International Cooperation Agency (JICA) (Watanabe, 2015). Je nach Zeitpunkt bewerteten die Autor:innen darin entweder den Beitrag der Entwicklungsmaßnahme zur Schaffung von Rahmenbedingungen für Nachhaltigkeit oder die tatsächliche Aufrechterhaltung von Kapazitäten über das Ende der Maßnahme hinaus. Im zweiten Teil thematisierte Alberti die Herausforderung bei der Anwendung des Kriteriums in der humanitären Hilfe. Einerseits verfolgten humanitäre Interventionen häufig einen kurzfristigen Ansatz, andererseits mache es fragile Staatlichkeit schwer, nachhaltigen Kapazitätsaufbau zu leisten. Dennoch werde Nachhaltigkeit auch in fragilen Kontexten evaluiert, beispielsweise durch die Prüfung der institutionellen Rahmenbedingungen für eine Weiterführung der Maßnahmen. Um die Herausforderungen der Evaluierung von Nachhaltigkeit in der humanitären Hilfe zu adressieren, erarbeitet das DAC EvalNet derzeit eine Handreichung, die eine Orientierung für unterschiedliche Kontexte geben soll.

Der nachfolgende Beitrag von Sarah Gharbi vom Active Learning Network for Accountability and Performance (ALNAP) knüpfte daran an und stellte Ergebnisse einer Studie zur Anwendung des Prüfkriteriums Nachhaltigkeit in Evaluierungen der humanitären Hilfe vor. Die Studienergebnisse dienten auch als Grundlage für die Überarbeitung der ALNAP-Handreichung von 2006 zur Anwendung der OECD DAC-Evaluierungskriterien in der humanitären Hilfe (ALNAP, 2006). Während in der Handreichung von 2006 die Überprüfung der *connectedness* (Anschlussfähigkeit der Maßnahme) im Vordergrund stand, wird das Kriterium in der überarbeiteten Handreichung neu definiert und als *inter-connection* aufgenommen (ALNAP, 2023). Dabei wird Anschlussfähigkeit in seiner zeitlichen Dimension um eine relationale Dimension erweitert, die die Bezüge von humanitären zu entwicklungspolitischen und friedensfördernden Interventionen entsprechend des Humanitarian-Development-Peace (HDP) Nexus Ansatzes explizit in die Bewertung einbezieht.

Martin Noltze vom Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval) thematisierte die unterschiedliche Anwendung des Prüfkriteriums in Evaluierungen in der deutschen EZ (Noltze et al., 2018). Die reformierten Orientierungslinien des Bundesministeriums für Wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) definieren analog zum OECD DAC Nachhaltigkeit als Dauerhaftigkeit der Wirkungen (BMZ, 2021). Zudem wird der Multidimensionalität von Nachhaltigkeit Rechnung getragen. Noltze zufolge gelte es, im Fall von humanitären Maßnahmen neben Anschlussfähigkeit auch Resilienz zu berücksichtigten. Die Evaluierung solle zudem den Zeithorizont der transformativen Veränderung berücksichtigen und eine intersektionale Perspektive einnehmen. Durch die Reformierung habe sich die Vergleichbarkeit von Evaluierungsergebnissen verbessert und dennoch nicht zu signifikant anderen Kriterienbewertungen geführt.

Im Schlussvortrag des Einführungsblocks referierten Kathrin Greve und Afra Herr über Nachhaltigkeit in Evaluierungen des Auswärtigen Amtes (AA). Der Fokus von Evaluierungen im AA liegt auf der Bewertung von Maßnahmen der humanitären Hil-

fe in fragilen Kontexten und von Stabilisierungsmaßnahmen. Dabei wird Nachhaltigkeit sowohl anhand des Nachhaltigkeitskriteriums des OECD DAC als auch anhand der Definition von ALNAP evaluiert. In ihrer Präsentation nahmen die Referentinnen Bezug auf den Evaluierungsbericht der GFA für das AA innerhalb der ressortgemeinsamen strategischen Evaluierung des zivilen Engagements Deutschlands in Afghanistan (Meiering et al., 2023). Dem Bericht zufolge verfolgte das AA in Krisensituationen mit seinem Stabilisierungsansatz das Ziel der Anschlussfähigkeit. Dabei sollten verfügbare Zeitfenster flexibel für schnelle und risikobehaftete Interventionen genutzt werden, um Veränderungen überhaupt erst anstoßen zu können, die langfristig mit anderen Politikbereichen verknüpft werden könnten (AA, 2022). Die erfolgreiche Implementierung in Krisenkontexten trage somit zu Nachhaltigkeit bei.

Außerdem betonten die Autor:innen die Bedeutung von Exit-Strategien, die bereits bei der Planung einer Intervention mitgedacht werden müssten. Der HDP-Nexus sei dabei ein wichtiger Ansatz, um die Brücke zwischen den unterschiedlichen Strategien der Ministerien zu schlagen. In der anschließenden Diskussion kam die Frage auf, wie das AA aus den Evaluierungsergebnissen lernen könnte. Greve erläuterte, dass Evaluierungsberichte des AA grundsätzlich nicht veröffentlicht, jedoch die Ergebnisse in internen Feedbackrunden zum institutionellen Lernen, etwa für die Strategieentwicklung und personelle Aus- und Weiterbildung genutzt würden.

Nachhaltigkeitsbewertung in der evaluativen Praxis

Der zweite Teil des ersten Tages begann mit einem World-Café, bei dem an drei Tischen das Nachhaltigkeitskriterium aus unterschiedlichen Perspektiven beleuchtet wurde. Der Austausch am ersten Tisch drehte sich um die Frage, wie Nachhaltigkeit in laufenden bzw. Schlussevaluierungen adäquat geprüft werden könne, z. B. wenn entweder noch keine Effekte sichtbar seien oder Messdaten fehlten. Im Hinblick auf Verbesserungspotentiale für die Anwendung des Kriteriums kamen die Teilnehmenden darin überein, das bereits bei der Projektplanung sowie bei der Entwicklung von Leistungsbeschreibungen angesetzt werden müsse. Darin solle reflektiert und festgelegt werden, wer von den Ergebnissen lernen und welcher Aspekt des Kriteriums besonders betrachtet werden solle. Auch sollten Indikatoren zur Messung schon bei der Projektplanung mitgedacht werden. Eine gute Hypothesenbildung über die beabsichtigten Wirkungen sei dabei von großer Bedeutung.

Am zweiten Tisch tauschten sich die Teilnehmenden darüber aus, auf welche Weise Nachhaltigkeit in all seinen Dimensionen in Evaluierungen bearbeitet wird. In der Reflexion kamen einige Teilnehmende zu dem Ergebnis, dass ein bewusst ermöglichter Interpretationsspielraum für das Nachhaltigkeitskriterium in Leistungsbeschreibungen Evaluator:innen eine flexiblere Fokussierung auf unterschiedliche Nachhaltigkeitsdimensionen ermögliche. Obwohl Nachhaltigkeit als OECD-DAC-Kriterium obligatorisch evaluiert werden sollte, wurde von Teilnehmenden positiv wahrgenommen, dass die Anwendung des Kriteriums von Auftraggebenden nicht immer eingefordert würde. Gleichzeitig berichteten viele Auftragnehmende von einem Erkenntnisgewinn über die Mehrdimensionalität von Nachhaltigkeit. Obwohl die unterschiedlichen Interpre-

tationen von Nachhaltigkeit als bereichernd empfunden wurden, berichteten Teilnehmende auch von Herausforderungen für die Messung. Gerade weil EZ einen transformativen Ansatz verfolge, sei es wichtig, Nachhaltigkeit im Sinne eines transformativen Wandels zu verstehen, statt nur auf die Dauerhaftigkeit von Wirkungen zu fokussieren. So wirkten beispielsweise biographische Veränderungen bei den Zielgruppen, initiiert etwa durch Bildungsprojekte, unabhängig von der Fortführung einer Maßnahme lebenslang nach. Auch die Abgrenzung des Nachhaltigkeits- vom Impact-Kriterium warf bei Teilnehmenden Fragen auf.

Der dritte Tisch thematisierte Nachhaltigkeit im Monitoring. Wichtig für das Monitoring sei vor allem die kontinuierliche Datensammlung ab Maßnahmenbeginn. Gleichzeitig könnten nicht alle Veränderungen, zum Beispiel im Verhalten von Zielgruppen, dauerhaft gemessen werden.

Anschießend an das World-Café präsentierten Birgit Hansl, Elena Bardasi, Stephan Wegner und Mees van der Werf von der Independent Evaluation Group (IEG) der Weltbankgruppe einen Nachhaltigkeitsansatz anhand von drei strategischen Evaluierungen. Das erste Fallbeispiel stellte eine Evaluierung zu Geschlechterungleichheiten in fragilen Konfliktkontexten dar. Im Zentrum der Maßnahme stand die ökonomische Stärkung von Frauen und die Prävention von geschlechterbasierter Gewalt. Nachhaltigkeit wurde darin definiert als die Nachhaltigkeit der Veränderungen auf der Ebene der Zielgruppe zum Zeitpunkt der Implementierung sowie explizit darüber hinaus. Diese könnte durch drei Ansätze erreicht werden: 1. eine geographische Ausweitung und Vergrößerung der Zielgruppe, 2. eine systematische Stärkung der gesellschaftlichen Umgebung von Frauen und Mädchen und 3. die Nutzung von Synergieeffekten durch die Zusammenarbeit mit anderen Geber:innen. Im Ergebnis zeigte sich, dass nur wenige Projekte Nachhaltigkeit in der Planung berücksichtigt hatten. Eine noch geringere Zahl zielte mit seinen Maßnahmen auf ein nachhaltig förderliches Umfeld für Frauen und Mädchen. In fragilen Konfliktkontexten setzten die meisten Projekte stattdessen lediglich auf die Fortführung und den Ausbau bestehender Maßnahmen. Notwendig sei jedoch ein Wandel hin zu einem strategischen Ansatz auf Länderebene (Weltbank, 2023).

Als zweites Beispiel präsentierte die IEG ihre Länderprogrammevaluierung Somalia. Zwischen 2013 und 2023 unterstützte die Weltbank den Aufbau staatlicher Institutionen und leistete einen Beitrag zur Gewährleistung der Daseinsvorsorge. In der Evaluierung wurde Nachhaltigkeit entlang von drei Dimensionen bewertet: finanzielle Nachhaltigkeit, Eigenverantwortung der Partnerorganisation und institutionelle Stabilität. Im fragilen Kontext Somalias zeigte sich die nachhaltige Implementierung aufgrund der schwachen staatlichen Strukturen besonders herausfordernd, da die Ausweitung der Daseinsvorsorge die finanzielle Belastung des Staates erhöhte, was sich negativ auf die finanzielle Nachhaltigkeit auswirkte (Weltbank, 2025).

Das dritte und letzte Beispiel thematisierte die Herausforderung bei der Evaluierung von Maßnahmen, die durch Drittparteien implementiert werden. Der IEG zufolge gehe dies mit einer erschwerten Evaluierbarkeit und verringerten Möglichkeiten des institutionellen Lernens einher.

Den Abschlussvortrag des ersten Tages hielt Lea Kleinsorg vom DEval zum Thema „Foresight trifft Evaluation: Integration von Foresight-Methoden in Evaluierungen". Darin thematisierte sie die Herausforderung, dass klassische Ex-Post-Evaluierungen

Geschehnisse der Vergangenheit evaluieren, um über ungewisse Zukünfte zu informieren. Die zunehmend komplexere Gegenwart und Zukunft erschwere es jedoch, aus rückblickenden Evaluierungen Empfehlungen abzuleiten. Um die Ungewissheit zu adressieren, würden durch die Anwendung von Foresight-Methoden kommende Entwicklungen bewusst im Plural gedacht und in vorausschauenden Evaluierungen untersucht. Methodisch ordneten sich solche Evaluierungsansätze der sogenannten fünften Generation von Evaluierungsansätzen zu (Beywl, 2024). Kernvoraussetzung sei das „Futures Thinking" (Gardner et al., 2024, S. 6), also die Fähigkeit und der Wille, sich abstrakt mit unsicheren Zukunftsszenarien zu beschäftigen. Die systematische Berücksichtigung von Unsicherheiten stärke die methodologische Qualität, partizipative Ansätze und die Theorien des Wandels in Foresight-Evaluierungen, um besser informierte Empfehlungen zu ermöglichen und Nachhaltigkeit in zukünftigen Interventionen sicherzustellen.

Werkstattgespräche zur Evaluierung von Nachhaltigkeit

Der zweite Tag begann mit einem Vortrag von Tilman Welte von der GFA Consulting Group, der Ergebnisse aus den ressortgemeinsamen Evaluierungen zu Irak (Hartmann et al., 2022) und Afghanistan (Hartmann et al., 2023) vorstellte. Welte ging auf die Fragilitäts- und Volatilitätsdefinitionen von EU und OECD ein und verwies auf die unterschiedlichen Ansprüche an Nachhaltigkeit an der Schnittstelle zwischen humanitärer Hilfe, Stabillisierung und langfristiger Entwicklungspolitik. Aufgrund unterschiedlicher Priorisierungen und politischer Zielsetzungen blieb die Zusammenarbeit zwischen AA und BMZ entlang des HDP-Nexus im Irak und Afghanistan beschränkt. Um Anschlussfähigkeit und Nachhaltigkeit auch in fragilen Kontexten zu erreichen, sei es notwendig, die interministerielle Zusammenarbeit zu verbessern sowie Exit-Strategien und Maßnahmen zur Anschlussfähigkeit und Verantwortungsübernahme durch Partner:innen bereits in der Planung mitzudenken.

Anschließend an den Vortrag verteilten sich die Teilnehmenden auf drei Werkstattgespräche. Das erste Gespräch befasste sich mit der Evaluierung und dem Monitoring von Nachhaltigkeit in fragilen Kontexten. Chiara Pizzuti und Berit Mohr von der GFA Consulting Group präsentierten Möglichkeiten des Einsatzes von Fernerkundungsdaten für die Bewertung von Projektwirkungen in fragilen Kontexten am Beispiel von Afghanistan und Äthiopien. Mithilfe von Fotos und Satellitenbildern konnte bereits während der Implementierung der Projektfortschritt gemessen werden (GFA, 2025). Vivien Mohrland von der KfW Entwicklungsbank ging in ihrem Vortrag auf das Spannungsfeld zwischen den kurzfristig ausgerichteten Projektzielen einerseits und dem Anspruch an nachhaltigen Wirkungen andererseits bei der Evaluierung von Projekten entlang des HDP-Nexus am Beispiel zweier Projektevaluierungen in Syrien ein. Dabei verwies sie auf die KfW-Handreichung zur Bewertung des Nachhaltigkeitskriteriums (KfW, 2024). Um die Nachhaltigkeit von humanitären Maßnahmen in fragilen Kontexten mit häufig schwachen Kapazitäten von Partner:innen und Zielgruppen kontextadäquat – d. h., mit begrenztem Anspruch – zu evaluieren, orientiert sich die KfW an einem mehrstufigen Verfahren. Dabei kann ein begrenzter Anspruch an

Nachhaltigkeit entweder aus Projektunterlagen identifiziert und durch die Evaluierungseinheit validiert, oder nachträglich durch die Evaluierungseinheit im Rahmen der Kontextanalyse bestimmt werden. Wird ein begrenzter Anspruch an Nachhaltigkeit festgestellt, wird das Nachhaltigkeitskriterium bei der Gesamtbewertung der Maßnahme nicht berücksichtigt.

Das zweite Werkstattgespräch nahm Nachhaltigkeit in formativen und Schlussevaluierungen aus der Perspektive von Auftraggebenden und Auftragnehmenden in den Fokus. Benjamin Frech von der Deutschen Gesellschaft für Internationale Zusammenarbeit (GIZ) präsentierte die Ergebnisse der jüngsten Auswertungen der GIZ-Stabsstelle Evaluierung aus der Perspektive der Auftraggebenden. Nachhaltigkeit wird darin in Anlehnung an die OECD DAC als Dauerhaftigkeit bzw. Absehbarkeit der Dauerhaftigkeit von Wirkungen auf der Outcome- und Impactebene insbesondere nach dem Förderende verstanden, unter Berücksichtigung eingetretener bzw. absehbarer Risiken. Die Auswertungen zeigten, dass Nachhaltigkeit und Wirkungen aufgrund des stärkeren Einflusses externer Faktoren schwerer zu steuern seien als Effizienz und Effektivität. Demnach seien oft nur Aussagen mit Prognosecharakter über die Plausibilität dauerhafter Wirkungen möglich. Eine längere Projektlaufzeit habe tendenziell einen positiven Einfluss auf das Erreichen dauerhafter Wirkungen. Gleichzeitig erschwerten oft fehlende Übergabe- und Ausstiegsstrategien die nachhaltige Wirksamkeit.

Regine Parkes (FAKT Consult) präsentierte die Perspektive der Auftragnehmenden. Die Untersuchung von Nachhaltigkeit auf Projektebene, z. B. in Zentralen Projektevaluierungen der GIZ sei bei summativen Evaluierungen nur dann möglich, wenn das Potenzial für die Dauerhaftigkeit der Projektziele erkenn- und belegbar sei – bei formativen Evaluierungen müssten zumindest die Grundlagen dessen erkenn- und belegbar sein. Auch Parkes betonte die Bedeutung von Exitstrategien bzw. Folgevorhaben. Über die Projektebene hinausgehend intendierte Wirkungen seien in Zentralen Projektevaluierungen grundsätzlich nicht messbar.

Das dritte Werkstattgespräch thematisierte die Multidimensionalität von Nachhaltigkeit in Evaluierungen. Sonja Kind stellte den von der Ad-hoc-Gruppe Nachhaltigkeit der DeGEval erarbeiteten Leitfaden zur Evaluierung von Nachhaltigkeit vor (Kerlen et al., 2025). Der Leitfaden richtet sich an Evaluierende, Auftraggebende und weitere Interessierte. Ziel der Ad-Hoc-Gruppe war es, das Verständnis des Nachhaltigkeitsbegriffs zu verbessern. In ihrer Nachhaltigkeitsdefinition orientierte sich die Ad-Hoc-Gruppe an der Definition des Rates für nachhaltige Entwicklung. Nachhaltigkeit wird darin entlang der ökonomischen, ökologischen und sozialen Dimension definiert und nimmt Zielkonflikte bei der Bedürfnisbefriedigung heutiger und zukünftiger Generationen in den Blick. Mit dieser komplexen Nachhaltigkeitsdefinition geht die Ad-hoc-Gruppe über den Nachhaltigkeitsbegriff des OECD DAC hinaus. Der Entwurf der Ad-hoc-Gruppe, der Evaluierungsleitfragen zu den unterschiedlichen Dimensionen enthält, wurde auf der Internetseite der DeGEval veröffentlicht (Kerlen et al., 2025). Der Leitfaden schließt mit der Empfehlung, Nachhaltigkeit eng auf den Evaluierungsgegenstand zu beziehen und zu messen.

Im zweiten Teil stellte Katharina Kreutz den multidimensionalen Evaluierungsansatz der Welthungerhilfe vor. Darin werden die unterschiedlichen Dimensionen von Nachhaltigkeit in ihrer Wechselwirkung und ihrem Einfluss auf die Dauerhaftigkeit der Wirkungen untersucht. Für die zielgerichtete Bewertung von Nachhaltigkeit wür-

den die Dimensionen im gesamten Projektzyklus der Welthungerhilfe systematisch mitgedacht. Gleichzeitig ermögliche der Ansatz eine Fokussierung auf ausgewählte, relevante Dimensionen bei der Kriteriumsbewertung.

Beendet wurde die Frühjahrstagung mit einer Fishbowl-Diskussion über die Mittelkürzungen in der EZ und Implikationen für Nachhaltigkeit und Evaluierung. Anja Desai von der GFA Consulting Group, Stefan Ewers von CARE, Benjamin Frech von der GIZ und Petra Kiel von der Christoffel Blindenmission gaben die Folgen der Haushaltssituation aus der Perspektive der unterschiedlichen Akteur:innen wieder. Besonders deutlich wurde, welche negativen Folgen die Kürzungen insbesondere auf zivilgesellschaftlicher Seite für die Zusammenarbeit mit vulnerablen Gruppen haben. Auch eine zunehmende politische Einflussnahme auf die Wortwahl in Projektbeschreibungen und Anträgen, etwa mit Bezug zu Gender, erschwere die Evaluierung, da thematische Stichwortsuchen deutlich erschwert würden. Wenn Institutionen geschlossen würden, gingen damit auch immer institutionelles Wissen und Daten verloren, die nicht mehr für Evaluierungen genutzt werden könnten.

Die Frühjahrstagung zeigte, dass Nachhaltigkeit obgleich ihrer Verankerung in zentralen Evaluierungsleitlinien heterogen definiert und bewertet wird. Während der Fokus in der humanitären Hilfe auf der Ermöglichung einer Anschlussfähigkeit an langfristig ausgerichtete Maßnahmen und der Stärkung der Resilienz von Rechteinhabenden liegt, zielt EZ auf eine langfristige Transformation von Lebensumständen. Die Debatten rund um Definitionen und Dimensionen zeigten, dass es für die Evaluierung von Nachhaltigkeit von großer Bedeutung ist, die unterschiedlichen Dimensionen und ihre Wechselwirkungen auf den verschiedenen Ebenen systematisch mitzudenken und kontextadäquat zu definieren.

Literatur

ALNAP. (2006). *Evaluating humanitarian action using the OECD-DAC criteria. An ALNAP guide for humanitarian agencies*. https://alnap.org/help-library/resources/evaluating-humanitarian-action-using-the-oecd-dac-criteria/

ALNAP. (2023). *Review of the OECD DAC criteria for evaluating humanitarian action*. https://alnap.org/help-library/resources/summary-brief-revie https://alnap.org/help-library/resources/summary-brief-review-of-the-oecd-dac-criteria-for-evaluating-humanitarian-action/w-of-the-oecd-dac-criteria-for-evaluating-humanitarian-action/

Auswärtiges Amt. (2022). *Stabilisierung gestalten. Außen- und sicherheitspolitisches Konzept für ein integriertes Friedensmanagement*. https://www.auswaertiges-amt.de/resource/blob/2571114/c88edabb5da7869a111d5a0264f2d0c0/integriertes-friedensengagement-data.pdf

Beywl, W. (2024, 18.–20. September). *«Evidenz» evaluationswissenschaftlich-transdisziplinär betrachtet* [Konferenzbeitrag]. 27. Jahrestagung der Deutschen Gesellschaft der Evaluation (DeGEval), Potsdam, Deutschland. https://www.degeval.org/fileadmin/content/Z04_Veranstaltungen/Jahrestagungen/Potsdam_2024/Dokumentation/Sessions_A/2024_A5_Beywl_Intro.pdf

BMZ. (2021). *Grundsätze der Evaluierung. Evaluierungskriterien der bilateralen Entwicklungszusammenarbeit. Orientierungslinien des BMZ*. https://www.bmz.de/resource/blob/92894/evaluierungskriterien.pdf.

Gardner, A., Bontoux, L., & Barela, E. (2024). *Applying Evaluation Thinking to Practice and Foresight*. Association of Professional Futurists. https://www.apf.org/_files/ugd/682b5a_dec6819cdcc742b4880d23ee04e89e70.pdf

GFA Consulting Group. (2025). *Digital Innovation Unit*. https://www.gfa-group.de/area/downloads/634556/GCG_FS_Digital_Innovation_E.pdf

Hartmann, C., Roxi, H., Atal, M. R., Berchthold, H., Kellogg, M., Weeger, M., Marquardt, U., Lange, A., Mielke, K., Seltmann, P., Vogt, T., Meiering, G., Welte, T., Scheuermann, P., Hennefeld, V., Cain, E., Keyl, L., Kuhn, S., Brinkmann, C., Sickert, M., & Steinitz, N. (2023). *Ressortgemeinsame strategische Evaluierung des zivilen Engagements der Bundesregierung in Afghanistan*. Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), Deutsche Hochschule der Polizei (DHPol), GFA Consulting Group. https://www.deval.org/fileadmin/Redaktion/PDF/05-Publikationen/Berichte/2023_Afghanistan/2023_DEval_Bericht_Afghanistan_ressortgemeinsam_web.pdf

Hartmann, C., Wedel, R., vom Hofe, M., Weeger, M., & Winterhoff, L. (2022). *Ressortgemeinsame strategische Evaluierung des AA- und des BMZ Engagements in Irak. Ressortspezifischer Bericht zum Engagement des BMZ in Irak*. Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval). https://www.bmz.de/resource/blob/106602/ressortspezifischer-bericht-bmz.pdf

ICAI. (2018). *Assessing DFID's results in improving Maternal Health. An impact review*. https://icai.independent.gov.uk/wp-content/uploads/ICAI-review-Assessing-DFIDs-results-in-improving-Maternal-Health-.pdf

Kerlen, C., Kind, S., Finkemeier, T., Jessing, D., Kofler, J., Nowak, C., Wieser, H., Wolf, B., & Zweidler, R. (2025). *Evaluationskriterium Nachhaltigkeit. Leitfaden für die Verwendung*. Ad-hoc-Gruppe Nachhaltigkeit, Gesellschaft für Evaluation (DeGEval). https://www.degeval.org/fileadmin/content/Z02_Arbeitskreise/Ad-hoc-Gruppen/Nachhaltigkeit/2025-05-07_Evaluationskriterium_Nachhaltigkeit.pdf

KfW. (2024). *Sustainability Guideline. Assessment and management of Environmental, Social, and Climate Aspects: Principles and Procedures*. https://www.kfw-entwicklungsbank.de/PDF/Download-Center/PDF-Dokumente-Richtlinien/Nachhaltigkeitsrichtlinie_EN.pdf

Meiering, G., Welte, T., Scheuermann, P., Hennefeld, V., Cain, E., Keyl, L., Kuhn, S., Brinkmann, C., Sickert, M., & Steinitz, N. (2023). *Finaler ressortspezifischer Evaluierungsbericht Afghanistan*. GFA Consulting Group, CEval GmbH, SiCon International Development GmbH. https://www.auswaertiges-amt.de/resource/blob/2635536/0620a09f0449fd743ef20a3184ccd6a5/231212-afg-ressortspezifisch-aa-data.pdf

Noltze, M., Euler, M., & Verspohl, I. (2018). *Nachhaltigkeit in der deutschen Entwicklungszusammenarbeit. Meta-Evaluierung*. Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval). https://www.deval.org/de/publikationen/meta-evaluierung-von-nachhaltigkeit-in-der-deutschen-entwicklungszusammenarbeit

OECD. (2019). *Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use*. OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2019/12/better-criteria-for-better-evaluation_f7a307eb/15a9c26b-en.pdf

OECD. (2021). *Applying Evaluation Criteria Thoughtfully*. OECD Publishing. https://www.oecd.org/en/publications/2021/03/applying-evaluation-criteria-thoughtfully_45a54ea7.html

Orth, M., Birsan, M., & Gotz, G. (2018). *The Future of Integrated Policy-Based Development Cooperation: Lessons from the Exit from General Budget Support in Malawi, Rwanda, Uganda and Zambia*. Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval). https://www.deval.org/fileadmin/Redaktion/PDF/05-Publikationen/Berichte/2018_Budget_Support/DEval_Executive_Summary_The_Future_of_Integrated_Policy-Based_Development_Cooperation_2018_EN.pdf

Watanabe, K. (2015). *FY 2015 Ex-Post Evaluation of Technical Cooperation Project "Inter-Communal Rural Development Project"*. Mitsubishi UFJ Research & Consulting. https://www2.jica.go.jp/en/evaluation/pdf/2015_0603847_4_f.pdf

Weltbank. (2023). *Addressing Gender Inequalities in Countries Affected by Fragility, Conflict, and Violence: An Evaluation of the World Bank Group's Support*. Independent Evaluation Group (IEG). https://ieg.worldbankgroup.org/sites/default/files/Data/Evaluation/files/gender-in-FCV.pdf

Weltbank. (2025). *The World Bank Group in Somalia. Fiscal Years 2013–23. Country Program Evaluation.* Independent Evaluation Group (IEG). http://documents.worldbank.org/curated/en/099307503252514921

Julian Friedrich | DEval – Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit | Fritz-Schäffer-Str. 26, D-53113 Bonn | E-Mail: Julian.Friedrich@DEval.org