

INDEED

Evidence – Based Model for Evaluation of
Radicalisation Prevention and Mitigation



This project has received funding by the European Union's
Horizon 2020 research and innovation programme
H2020-SU-SEC-2020 under grant agreement no 101021701

INDEED E-GUIDEBOOK 1

EVIDENZBASIERTE EVALUIERUNG VON PVE/CVE- UND DE- RADIALISIERUNGSINITIATIVEN

Grundsätze, Herausforderungen und Methoden

Leena Malkki, Mina Prokic und Irina van der Vet

INHALTSÜBERSICHT

DANKSAGUNGEN	4
EINFÜHRUNG	4
1. PVE/CVE/DERAD-INITIATIVEN	6
2. WAS IST EVIDENZBASIERTE EVALUIERUNG?	8
3. WARUM EVALUIEREN?	11
4. WANN SIND EVALUIERUNGEN ZU PLANEN UND DURCHZUFÜHREN?	12
5. WER SOLLTE EVALUIEREN?	13
5.1 Auswahl des:r Evaluators:in	13
5.2 Einbindung von Stakeholdern	14
6. HAUPTARTEN DER EVALUATION	16
6.1 Formative Evaluierung	16
6.2 Prozessevaluierung	18
6.3 Ergebnisevaluierung	19
6.4 Andere Evaluierungsarten	21
6.5 Kombination von Evaluierungsarten	23
7. EVALUIERUNGSKONZEPTE UND -METHODEN	24
7.1 Evaluierungsentwürfe	24
7.2 Methoden der Datenerhebung	28
7.3 Methoden der Datenanalyse	32
8. MENSCHENRECHTSBASIERTER ANSATZ DER EVALUIERUNG	34
9. ACHT TIPPS FÜR ERFOLGREICHE EVIDENZBASIERTE EVALUATIONEN	37
QUELLEN UND LITERATUR	40

DANKSAGUNGEN


Das INDEED E-Guidebook 1 wurde von Leena Malkki, Mina Prokic und Irina van der Vet (Universität Helsinki) verfasst.

Diese Arbeit wurde von anderen Projektpartner:innen unterstützt, insbesondere von Stephan Klose (Vrije Universiteit Brussel), der an der Abfassung der Kapitel 5 und 7 beteiligt war und eine Schlüsselrolle in der Entwicklung des Konzepts des evidenzbasierten Ansatzes für die Evaluierung spielte, Marzena Kordaczuk-Wajs (Polish Platform for Homeland Security), die an der Abfassung des Kapitels 6.1 beteiligt war sowie von Hannah Reiter und Norbert Leonhardmair (VICESSE), die zu Kapitel 8 beitrugen.

Die Autorinnen möchten sich bei den externen Gutachter:innen für ihre sehr nützlichen Kommentare zu früheren Versionen der E-Guides bedanken: Markus Kaakinen (Universität Helsinki), Joel Busher (Universität Coventry), Ioan Durnescu (RAN-Rehabilitationsgruppe), Katrien Van Mele (Verband der flämischen Städte und Gemeinden), Marteen Dewaele (Verband der flämischen Städte und Gemeinden), Anneli Portman (Finnisches Institut für Gesundheit und Wohlfahrt) und Miila Lukkarinen (Stadt Helsinki).

EINFÜHRUNG

Die INDEED E-Guidebooks sollen ein **gutes Verständnis des evidenzbasierten Evaluierungsansatzes vermitteln und aufzeigen, wie dieser im Bereich der Prävention und Bekämpfung von gewalttätigem Extremismus bzw. der Unterstützung von Deradikalisierung (PVE/CVE/Derad) angewandt werden kann**, und eine Orientierungshilfe für die Gestaltung solcher Initiativen bieten. Die E-Guidebooks richten sich hauptsächlich an **Praktiker:innen und politische Entscheidungsträger:innen**, die in diesem Bereich tätig sind. Sie können auch für professionelle Evaluator:innen und Akademiker:innen nützlich sein, die an der Evaluierung solcher Initiativen beteiligt sind und sich mit der evidenzbasierten Evaluierung oder der Durchführung von Evaluierungen speziell im Bereich PVE/CVE/Derad vertraut machen wollen.



INDEED E-Guidebook 1 bietet eine knappe Einführung in die Grundlagen der Evaluierung und in die Bedeutung eines evidenzbasierten Evaluierungsansatzes. Es stellt verschiedene Evaluierungsarten, -designs und -methoden vor und gibt Hinweise, wann und von wem Evaluierungen durchgeführt werden sollten. Es enthält ein Kapitel über Ethik und Evaluierung sowie zahlreiche Beispiele für frühere Evaluierungen von PVE/CVE/Derad-Initiativen.

INDEED E-Guidebook 2 vertieft die Anwendung eines evidenzbasierten Ansatzes für die Evaluierung. Es bietet Information, wie Evaluierung in die Konzeption von Initiativen integriert werden kann, sowie die schrittweise Anleitung für die Konzeption und Durchführung einer evidenzbasierten Evaluierung. Außerdem enthält es weitere Informationen darüber, wie sichergestellt werden kann, dass die Evaluierung nach ethischen Gesichtspunkten durchgeführt wird.

Der Hauptzweck der INDEED E-Guidebooks besteht darin, Praktiker:innen und politische Entscheidungsträger:innen mit den Grundsätzen und Praktiken evidenzbasierter Evaluierung vertraut zu machen, damit sie als **gut informierte Beteiligte** an Evaluierungen agieren können und wissen, wie sie PVE/CVE/Derad-Initiativen so planen und umsetzen, dass sie effektiv evaluiert werden können. Um als **Evaluator:in** zu agieren, ist **ein umfassenderes Fachwissen** über Evaluierungspraktiken, -konzepte und -methoden **erforderlich**, als in den E-Guidebooks enthalten ist. Dies sollte bei der Entscheidung, wer als Evaluator:in fungieren soll, berücksichtigt werden.

Die E-Guidebooks sind **Teil des INDEED-Toolkits für Praktiker:innen und politische Entscheidungsträger:innen, das im Rahmen des von der EU finanzierten H2020-Projekts INDEED (2021-2024) entwickelt wurde**. Das Ziel des INDEED-Projekts war es, das Wissen, die Fähigkeiten und die Fertigkeiten von PVE/CVE/Derad-Praktiker:innen und politischen Entscheidungsträger:innen in Bezug auf evidenzbasierte Evaluierungen und Praktiken zu stärken. Dieses Toolkit enthält auch ein **INDEED-Evaluierungstool**, das detailliertere Tipps und Empfehlungen enthält. Das INDEED-Toolkit finden Sie unter www.indeedproject.eu.

1.PVE/CVE/DERAD-INITIATIVEN

In den letzten zwei Jahrzehnten haben viele Länder **Mittel und Wege** entwickelt, **um Radikalisierungsprozesse, die in gewalttätigen Extremismus münden, zu verhindern und zu bekämpfen und um De-Radikalisierung** und den Ausstieg aus dem Extremismus zu fördern. Derartige Initiativen sind vor allem in westeuropäischen Ländern verbreitet. Dies hat zur Einführung einer großen Vielfalt von Policies, Strategien und Praktiken geführt, die in diesem E-Guidebook als **PVE/CVE/Derad-Initiativen bezeichnet werden**.

Die PVE/CVE/Derad-Initiativen umfassen eine breite Palette von Aktivitäten, die sich in vielerlei Hinsicht stark unterscheiden.

An der Umsetzung können verschiedene Arten von **Akteur:innen beteiligt** sein, z. B. Nichtregierungsorganisationen, Praktiker:innen (Pädagog:innen, Gesundheits- und Sozialarbeiter:innen), Sicherheitsbehörden (Polizist:innen und Vollzugspersonal), Religionsgemeinschaften sowie Forscher:innen und andere Expert:innen. Es ist üblich und empfehlenswert, dass solche Initiativen ein gewisses Maß an behördenübergreifender Zusammenarbeit beinhalten.

Die Initiativen können in verschiedenen **sozialen Umfeldern stattfinden**, z. B. in Schulen, Gemeinde- und Religionszentren, Jugendzentren, Sozialdiensten, Sportvereinen oder Gefängnissen.

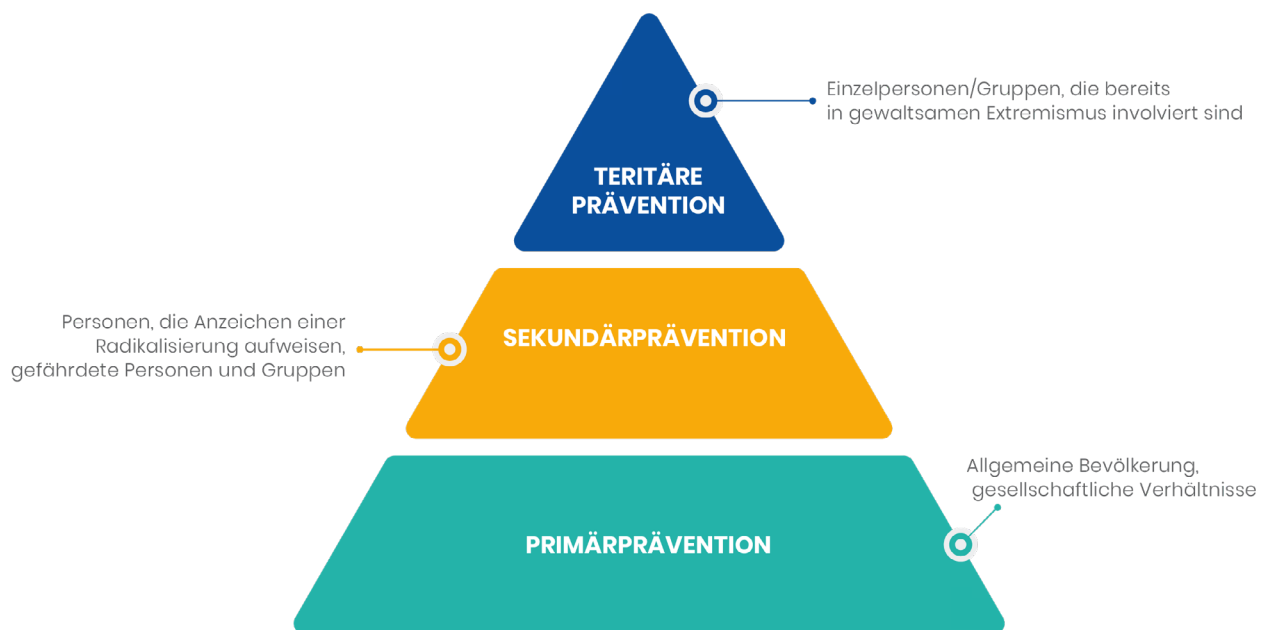
Auch der **Umfang und die Dauer** variieren. Einige Initiativen sind klein, während andere landesweit angelegt sind. Einige Initiativen sind langfristig angelegt oder sogar Teil ständiger täglicher Aktivitäten, während andere als kurzfristige Maßnahmen konzipiert sind. In diesem Zusammenhang gibt es auch Unterschiede bei der Finanzierung. Es ist üblich, dass Initiativen eine relativ kurzfristige Projektfinanzierung durch den Staat, die Kommunen, Wohlfahrtsverbände oder Stiftungen erhalten, während es einigen Initiativen gelungen ist, eine langfristige Finanzierung zu sichern. Die Initiativen unterscheiden sich darin, wer **Initiator:in** ist. Bei einigen handelt es sich um Top-Down-Initiativen, die vom Staat oder der Kommune entwickelt wurden, während andere Initiativen von einer NGO oder einer Community an der Basis entwickelt wurden.

Die Initiativen unterscheiden sich in ihrem **Schwerpunkt**. Einige Initiativen konzentrieren sich hauptsächlich darauf, die extremistischen Überzeugungen der Zielgruppe zu bekämpfen, während andere stattdessen darauf abzielen, die Beteiligung der Zielgruppe an gewalttätigen extremistischen Aktionen zu verhindern oder zu beenden. Viele Initiativen befassen sich bis zu einem gewissen Grad mit beiden Aspekten.

Es gibt erhebliche Unterschiede in Bezug auf die **Zielgruppen** und das **Stadium der Radikalisierung**, auf das sich die Initiativen konzentrieren. Es hat sich eingebürgert, den Bereich PVE/CVE/Derad in Anlehnung an das Modell der öffentlichen Gesundheit in drei Bereiche zu unterteilen:

- **Die Primärprävention** richtet sich an die gesamte Gesellschaft und zielt in erster Linie darauf ab, das Bewusstsein für Radikalisierung und gewalttätigen Extremismus zu schärfen und die Widerstandskraft gegen Radikalisierung zu stärken. Diese Initiativen haben in der Regel die Form von Programmen zur Förderung der Partizipation, zur Stärkung des Gefühls der Integration oder zur Vermittlung von Kultur- und Medienkompetenz. Sie können sich auch mit der sozialen Lage und Ungleichheit befassen, die als Radikalisierungsfaktoren gelten.
- **Die Sekundärprävention** befasst sich mit Einzelpersonen oder Gruppen, die bereits Anzeichen einer Radikalisierung aufweisen oder als radikalierungsgefährdet eingestuft werden. Zu den Maßnahmen gehören Mentoring, die Verbesserung der Beziehungen und der Kommunikation zwischen Behörden und gefährdeten Gruppen sowie der Aufbau von Resilienz in diesen Gruppen, die als gefährdet gelten.
- **Die tertiäre Prävention** zielt auf Einzelpersonen oder Gruppen ab, die bereits in gewalttätigen Extremismus verwickelt sind. Ziel ist es, sie zu ermutigen, sich von extremistischen Ideen und/oder Verhaltensweisen abzuwenden. Bei den Initiativen auf dieser Ebene handelt es sich in der Regel um individuelle Mentoring-Programme, die in unterschiedlichen Kontexten (Gefängnis, Bewährungshilfe, Gemeinde) und von verschiedenen Akteur:innen (Regierung, NRO) durchgeführt werden

Abbildung 1: Bereiche der PVE/CVE/Derad-Aktivitäten und ihre Zielgruppen



Mehr erfahren

- Busher, J. Malkki, L. & Marsden, S. (Hrsg.) (2023). Routledge Handbook on Radicalisation and Countering Radicalisation. Routledge.
- Köhler, D. (2017). Understanding Deradicalization: Methods, Tools and Programs for Countering Violent Extremism. Routledge.
- Byrne-Diakun, R.M. (Hrsg.). (2016). [Countering Violent Extremism: Applying the Public Health Model](#). Georgetown University, Zentrum für Sicherheitsstudien.
- [Introductory Guide: Countering Violent Extremism](#). CREST: Centre for Research and Evidence on Security Threats...

2. WAS IST EVIDENZBASIERTE EVALUIERUNG?

Unter Evaluierung versteht man im Allgemeinen eine möglichst systematische und unparteiische Evaluierung der Umsetzung oder der Auswirkungen einer Initiative. Sie wird in der Regel durchgeführt, um zu verstehen:

- Wie läuft die Umsetzung der Initiative?
- Welche Art von Ergebnissen hat die Initiative hervorgebracht?
- Stützt sich die Initiative auf solide Annahmen?

Evaluierung ist nicht dasselbe wie Monitoring, auch wenn diese Begriffe manchmal synonym verwendet werden. Auch ist Evaluierung nicht dasselbe wie Risikobewertung. In Textbox 1 wird erläutert, wie diese und andere verwandte Begriffe in diesem E-Guidebook verstanden werden und wie sie sich von der Evaluierung unterscheiden.



1 EVALUIERUNG UND ANDERE VERWANDTE BEGRIFFE

Evaluierung bedeutet, dass die Initiative so systematisch und unparteiisch wie möglich beurteilt wird. Sie kann zum Beispiel die Ergebnisse, die Wirksamkeit, den Erfolg der Umsetzung oder die zugrunde liegenden Annahmen analysieren.

Unter Monitoring versteht man die kontinuierliche und regelmäßige Erfassung von Daten während der gesamten Initiative. Der Zweck des Monitorings besteht darin, den Fortschritt zu dokumentieren und zu verfolgen. Die Daten werden in der Regel von denjenigen erhoben, die die Initiative umsetzen. Monitoring selbst beinhaltet keine Evaluierung, aber die während des Monitorings gesammelten Daten können oft als Daten für die Evaluierung verwendet werden.

Der Begriff "Risikobewertung" wird im Bereich PVE/CVE/Derad am häufigsten für die Bewertung des Radikalisierungsrisikos einer Person verwendet. Es wurden mehrere Risikobewertungsinstrumente entwickelt, die den Praktiker:innen bei der Durchführung solcher Bewertungen helfen und somit die Entscheidung über die beste Vorgehensweise in einer bestimmten Situation erleichtern.

Die Folgenabschätzung ist ein Prozess, der dazu dient, die möglichen Folgen und Auswirkungen eines Vorhabens zu ermitteln und zu verstehen. Sie kann durchgeführt werden, um Entscheidungsträger über mögliche negative Auswirkungen zu informieren oder Möglichkeiten zur Verbesserung der positiven Auswirkungen aufzuzeigen. Folgenabschätzungen werden in der Regel durchgeführt, wenn neue Gesetze oder politische Maßnahmen geplant oder später überarbeitet werden, um einen besseren Einblick in die Art der Auswirkungen zu erhalten (z. B. auf die Gleichstellung der Geschlechter, die Umwelt oder Kinder).

Dieses E-Guidebook basiert auf den Grundsätzen der **evidenzbasierten Evaluation (EBE)**.

2 INDEED-DEFINITION DER EVIDENZBASIERTEN EVALUATION (EBE)

Evidenzbasierte Evaluation ist "ein Prozess der Planung und Durchführung von Evaluationen, der verfügbare externe Evidenz, professionelles Fachwissen und die Werte, Präferenzen und Umstände der Beteiligten integriert" (INDEED D1.2).



Evidenzbasierte Evaluation ist im Wesentlichen ein Ansatz zur Planung und Durchführung von Evaluationen, der auf den Prinzipien der evidenzbasierten Praxis aufbaut:



Principles of evidence-based practice

Evidenz – Evaluierung wird unter Verwendung von Wissen über Evaluierungspraktiken und -methoden geplant und durchgeführt. Darüber hinaus beinhaltet sie die Analyse (und oft auch die Erhebung) von qualitativ hochwertigen empirischen Daten.

Stakeholder – Die Evaluation berücksichtigt den Kontext und die Werte, Bedürfnisse, Präferenzen und Umstände der wichtigsten Stakeholder. Sie zielt darauf ab, das Lernen und die Entwicklung der evaluierten Initiative oder, allgemeiner, des Bereichs PVE/CVE/Deradikalisierung zu unterstützen.

(Professionelle) Analyse – Der/Die Evaluator:in verfügt über ausreichende Kenntnisse sowohl über evidenzbasierte Evaluierungsverfahren als auch über PVE/CVE/Derad-Initiativen, um eine gut konzipierte Evaluierung durchzuführen und fundierte Schlussfolgerungen auf der Grundlage einer systematischen Analyse der Daten zu ziehen. Der/die Evaluator:in ist außerdem in der Lage, die Evaluierung unparteiisch und ethisch einwandfrei durchzuführen.

Eine evidenzbasierte Evaluierung unterscheidet sich von einem meinungsbasierten Evaluierungsprozess, der eher von Konventionen oder Intuition als von einer gründlichen Konsultation einschlägiger Forschungsarbeiten zu Evaluierungsdesigns oder einer systematischen Datenerhebung und -analyse geleitet wird. Sie unterscheidet sich auch von einem starren Evaluierungsprozess, der ohne angemessene Berücksichtigung der Präferenzen der Beteiligten oder des Kontexts und der Merkmale der untersuchten Maßnahme geplant und durchgeführt wird. Auf der Grundlage dieser Charakterisierungen lässt sich das Konzept der EBE grafisch in einer Matrix von vier Idealtypen verorten:

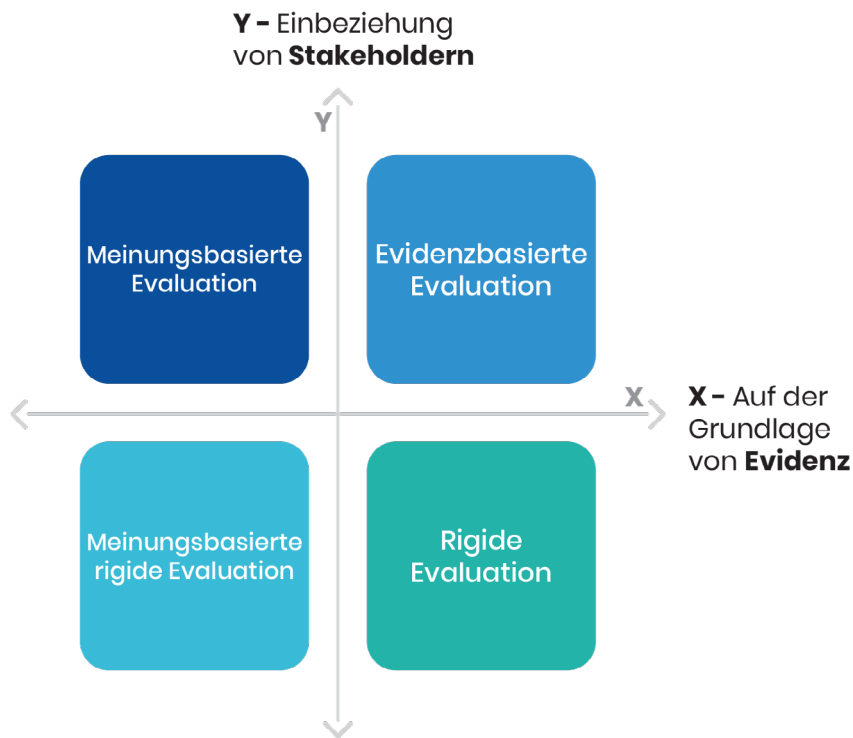


Abbildung 2: Vier Idealtypen der Evaluierung

Wenn Evaluierung nach dem evidenzbasierten Ansatz konzipiert und durchgeführt wird, erhöht sich ihre Qualität und ihr Nutzen. Die aktive Einbeziehung der Interessengruppen trägt dazu bei, dass die Evaluierung nützlich und relevant ist, und macht es wahrscheinlicher, dass ihre Ergebnisse genutzt werden. Die sorgfältige Berücksichtigung von Fakten bedeutet, dass die Evaluierung auf dem bestmöglichen Wissensstand beruht und relevante, repräsentative und zuverlässige Daten verwendet werden. Schließlich wird durch die Beachtung der fachlichen Kompetenz hervorgehoben, wie wichtig es ist, dass insbesondere der/die Evaluator:in, aber auch die wichtigsten beteiligten Akteur:innen über ein gutes Verständnis der Evaluierungspraktiken, der Initiative und ihres Kontexts verfügen.

Mehr erfahren

- INDEED-Deliverable D2.1 zu den Grundlagen der evidenzbasierten Praxis und der evidenzbasierten Evaluierung.
- INDEED E-Guidebook 2 – Gestaltung von PVE/CVE- und De-Radikalisierungsinitiativen und Evaluierungen nach den Grundsätzen der evidenzbasierten Praxis

3. WARUM EVALUIEREN?

Evaluierungen können eine wichtige Rolle bei der Entwicklung erfolgreicher und wirksamer Initiativen spielen. Die Evaluierung einzelner Initiativen wird auch die Grundlagen des Bereichs PVE/CVE/Derad im Allgemeinen stärken.

Evaluierungen unterstützen Lernen und Entwicklung

Die wertvollste Aufgabe von Evaluierungen ist es, den an der Initiative Beteiligten zu helfen, aus ihren Erfolgen und Misserfolgen zu lernen. Dies führt zu besser informierten Entscheidungen und schließlich zu besseren Leistungen. Eine Evaluierung ist am nützlichsten, wenn ihre Ergebnisse von denjenigen genutzt werden können, die die Initiative durchführen, und nicht mit Finanzierungsentscheidungen verbunden sind, die zwangsläufig einen Erfolgsdruck erzeugen.

Evaluierungen belegen den Beitrag einer Initiative

Evaluierungen können Praktiker:innen dabei helfen, ihren Kooperationspartner:innen und Geldgeber:innen die Schlüsselfaktoren zu erläutern, die ihre Arbeit erfolgreich machen, und dadurch anzugeben, welche Art von Ressourcen und Kapazitäten für die Fortsetzung ihrer Aktivitäten benötigt werden. Sie können auch die Praktiken, Werte und den spezifischen Beitrag einer Initiative zum Bereich PVE/CVE/Derad hervorheben.

Evaluierungen erhöhen die Transparenz und die öffentliche Rechenschaftspflicht

Viele PVE/CVE/Derad-Initiativen werden von einer öffentlichen Einrichtung, einer Wohltätigkeitsorganisation oder einer Stiftung finanziert. Es ist wichtig, nachweisen zu können, wofür die Mittel verwendet wurden, und sicherzustellen, dass sie wirksam eingesetzt wurden. Evaluierungen können diese Art von Informationen liefern und damit die Transparenz und öffentliche Rechenschaftspflicht einer Initiative erhöhen. Eine Evaluierung führt nicht nur zur öffentlichen Rechenschaftspflicht, weil sie dokumentiert, was erreicht wurde, sondern auch, weil sie Informationen darüber liefert, wie die Initiative verbessert werden kann. Auf diese Weise kann sie dazu beitragen, dass die Ressourcen in Zukunft noch besser genutzt werden.

Evaluierungen tragen dazu bei, eine solidere Faktengrundlage für PVE/CVE/Derad-Initiativen zu schaffen

Eine faktengestützte Evaluierung ist ein unverzichtbarer Bestandteil der Entwicklung evidenzbasierter Praktiken – mit anderen Worten: Initiativen, die auf der Grundlage der besten verfügbaren Erkenntnisse über Radikalisierung und deren Verhinderung konzipiert werden. Gegenwärtig sind die wissenschaftlichen Erkenntnisse über Radikalisierung und ihre Prävention noch begrenzt. Es ist immer noch unklar, was unter welchen Bedingungen im Bereich PVE/CVE/Derad funktioniert. Evaluierungen werden uns dabei helfen, mehr darüber zu erfahren, insbesondere wenn die Ergebnisse veröffentlicht werden, damit Forscher:innen, politische Entscheidungsträger:innen und Praktiker:innen sie für ihre Arbeit nutzen können.

4. WANN SIND EVALUIERUNGEN ZU PLANEN UND DURCHZUFÜHREN?

Es ist unmöglich, zu früh mit einer Evaluierung zu beginnen. Idealerweise sollte der Evaluierungsplan zusammen mit der Initiative selbst entwickelt werden.

Der beste Zeitpunkt, um sich Gedanken darüber zu machen, wann und wie die Initiative evaluiert werden soll, ist in der Planungsphase. Nur so ist es möglich, Datenerhebung und Monitoring so zu gestalten, dass alle erforderlichen Evidenzen zur Verfügung stehen, wenn es Zeit für die Evaluierung ist. Es kann möglich sein, Evaluierungen durchzuführen, die ursprünglich nicht geplant waren, aber es wird viele weitere Einschränkungen geben, was genau und wie evaluiert werden kann (siehe INDEED E-Guidebook 2 für weitere Informationen).



TIPP:

Idealerweise werden Evaluierungen als Teil der Konzeption der Initiative geplant.

Ein weiterer Grund, sich bereits in der Planungsphase der Initiative Gedanken über die Evaluierung zu machen, ist die Tatsache, dass Evaluierungen von den frühen Phasen der Initiative bis zu ihrem Ende durchgeführt werden können.

In der Planungsphase einer Initiative ist es möglich, den Umsetzungsplan und die zugrunde liegenden Annahmen zu evaluieren. Diese Art der Evaluierung wird als formative Evaluierung bezeichnet. Wenn die Initiative zunächst als Pilotprojekt erprobt wird, ist es möglich, das Pilotprojekt zu evaluieren. Diese Art der Evaluierung, die vor der Verabschiedung der endgültigen Fassung der Initiative durchgeführt wird, wird manchmal als Ex-ante-Evaluierung bezeichnet.

Während der Durchführung der Initiative kann evaluiert werden, wie die Umsetzung verläuft und ob die Initiative wie beabsichtigt funktioniert. Diese Art der Evaluierung wird als Prozessevaluierung bezeichnet. Evaluierungen können auf Schwachstellen hinweisen und aufzeigen, wie die Initiative verbessert werden kann. Wenn die Umsetzung der Initiative schon einige Zeit läuft, ist es auch möglich, einen ersten Blick auf die kurzfristigen Ergebnisse zu werfen.

Nach Beendigung der Initiative kann mit Hilfe einer Evaluierung festgestellt werden, ob die gewünschten Ziele erreicht wurden und wie dauerhaft die Wirkung war. Diese Art der Evaluierung wird als ErgebnisEvaluierung bezeichnet. Sie kann bei der Planung neuer Initiativen hilfreich sein und einen wichtigen Beitrag zu dem noch recht begrenzten Wissen über die Wirksamkeit von PVE/CVE/Derad-Initiativen leisten.



TIPP:

Im INDEED E-Guidebook 2 finden Sie Anleitungen, wie Sie eine Evaluation von Anfang an in die Planung der Initiative einbeziehen können.

5. WER SOLLTE EVALUIEREN?

Evaluierungen erfordern die Beteiligung einer Reihe relevanter Personen und eine Definition ihrer Rollen. Natürlich braucht jede Evaluierung eine:n Evaluator:in, aber auch eine Reihe von Stakeholder sollte aktiv beteiligt sein.

Ein evidenzbasierter Ansatz für Evaluierung erfordert eine sorgfältige Auswahl der:s Evaluators:in. Es ist auch wichtig, zu entscheiden und zu spezifizieren, wie genau die Beteiligten einbezogen werden, wie ihre Beziehung zum:r Evaluator:in aussieht und wie sie alle zusammenarbeiten werden.

5.1 AUSWAHL DES:R EVALUATORS:IN

Eine Evaluierung kann entweder von einem:r externen Evaluator:in oder einem:r internen Evaluator:in durchgeführt werden. Ein:e Evaluator:in kann eine einzelne Person sein, es kann aber auch ein Team von Evaluierungsexpert:innen sein.

Ein:e externe:r Evaluator:in ist eine Person, die nicht an der Initiative beteiligt ist oder eine wesentliche Beziehung zu ihr unterhält. Externe Evaluator:innen sind in der Regel Berater:innen oder akademische Forscher:innen.

Ein interner Evaluator:in ist jemand, der gegenwärtig Teil der Initiative oder der für sie verantwortlichen Organisation/Institution ist.

Ob ein:e externe:r oder ein:e interne:r Evaluator:in vorzuziehen ist, hängt von der jeweiligen Situation ab. Die wichtigsten Punkte, die berücksichtigt werden sollten, sind die folgenden:

Tabelle 1: Faktoren, die bei der Auswahl eines:r Evaluators:in zu berücksichtigen sind

Fachwissen im Bereich Evaluierung	Der/die Evaluator:in sollte über gute Kenntnisse der Evaluationskonzepte und der einschlägigen Methoden verfügen. Frühere Erfahrungen mit der Durchführung von Evaluierungen sind von Vorteil.
Kenntnisse im Bereich PVE/CVE/ Derad	Der/die Evaluator:in sollte über ein gutes Verständnis von PVE/CVE/ Derad als Politikbereich und der Arten von Initiativen verfügen. Der/die Evaluator:in sollte sowohl mit dem Kontext, in dem die
Kenntnis der Initiative und ihres Kontextes	Initiative durchgeführt wird, als auch mit der Initiative selbst vertraut sein.
Unparteilichkeit und Interessenkonflikte	Der/die Evaluator:in sollte in der Lage sein, die Initiative auf unparteiische Weise zu analysieren und somit keinen Interessenkonflikt mit der Evaluation oder der Initiative haben.

Zugang zu den Daten	Der/die Evaluator:in sollte Zugang zu allen für die Evaluation erforderlichen Daten haben (oder in der Lage sein, diese zu sammeln).
Kosten/Ressourcen und Verfügbarkeit	Der/die Evaluator:in sollte über ausreichende Ressourcen und Zeit verfügen, um seine Aufgabe zu erfüllen.
Vertrauen	Der/die Evaluator:in sollte in der Lage sein, das Vertrauen derjenigen Personen zu gewinnen, deren Mitarbeit für die Durchführung der Evaluation erforderlich ist.
Inanspruchnahme	Der/die Evaluator:in sollte über Glaubwürdigkeit und Kommunikationsfähigkeiten verfügen, die den Weg für die Nutzung der Ergebnisse ebnen.

Traditionell wird der Einsatz eines:r externen Evaluator:in bevorzugt, weil man glaubt, dass die Evaluation dadurch unparteiischer und unvoreingenommener wird. Es gibt jedoch auch Überlegungen, die für eine:n interne:n Evaluator:in sprechen. Zum Beispiel kennt ein:e interner Evaluator:in die Initiative und den Kontext oft besser, hat möglicherweise besseren Zugang zu Daten und kann besser Vertrauen bei den wichtigsten Interessengruppen aufbauen. Es wird auch angenommen, dass die Evaluationsergebnisse eher genutzt werden, wenn die Evaluation von einem:r internen Evaluator:in durchgeführt wird.

In der Regel wird ein:e externe:r Evaluator:in bevorzugt, wenn das Ziel darin besteht, die Wirksamkeit einer Initiative zu bewerten (Ergebnisevaluierung) und/oder wenn die Ergebnisse der Evaluierung erhebliche Auswirkungen auf die Zukunft der Initiative haben werden. Wenn der Zweck der Evaluation darin besteht, zu verstehen, wie die Initiative umgesetzt wurde (Prozessevaluierung) und/oder wenn sie zu Lernzwecken durchgeführt wird, kann auch eine interne Evaluation eine Option sein.

5.2. EINBINDUNG VON STAKEHOLDERN

Eine evidenzbasierte Evaluierung erfordert die sorgfältige Einbeziehung der Bedürfnisse, Werte und Umstände der Beteiligten in jeder Phase des Evaluierungsprozesses. Definitionsgemäß fördert sie daher einen an den Interessensgruppen orientierten Ansatz für die Evaluierung. Ein solcher Ansatz kann wiederum verschiedene Formen annehmen, einschließlich einer kollaborativen, partizipativen oder Empowerment-Evaluierung.

Bei einer **kollaborativen Evaluation** arbeitet der/die Evaluator:in während des gesamten Evaluationsprozesses kontinuierlich mit den Interessensgruppen zusammen, wobei er die Verantwortung für das Verfahren behält. Eine solche kollaborative Evaluierung kann dem:r Evaluator:in helfen, die Bedürfnisse, Werte und Umstände der Interessensgruppen in den verschiedenen Phasen des Evaluierungsprozesses, einschließlich der Vorbereitungs-, Konzeptions-, Durchführungs- und Nutzungsphasen, besser zu verstehen und darauf zu reagieren.

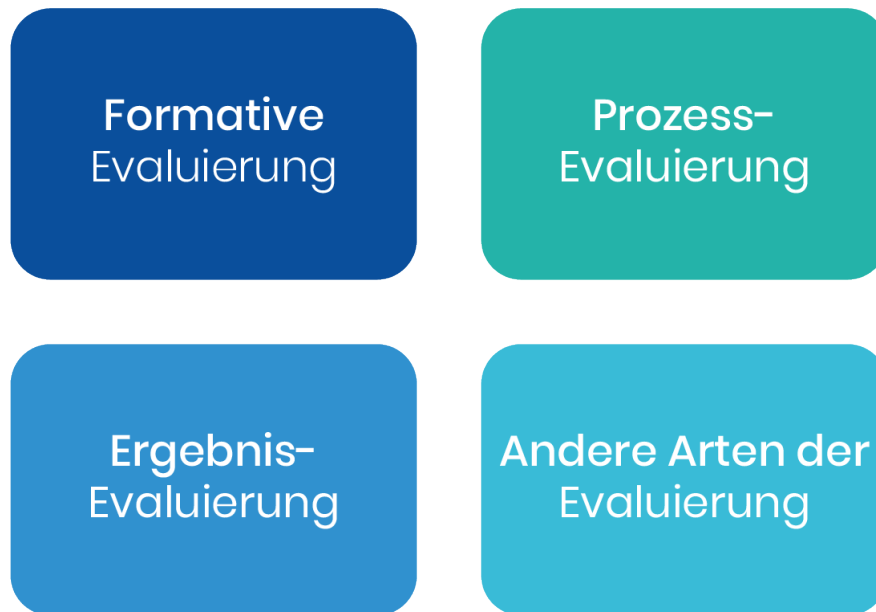
Bei einer **partizipativen Evaluation** teilen sich Evaluator:in und Stakeholder die Kontrolle über den Evaluationsprozess. Ein solcher Ansatz kann die gemeinsame

Entscheidungsfindung über die Evaluierungsziele, die Gestaltung und die Datenerfassung sowie die gemeinsame Durchführung und Nutzung einer Evaluation beinhalten. Dieser Prozess ermöglicht es den Beteiligten nicht nur, ihre Bedürfnisse und Werte zu äußern, sondern diese auch aktiv in den Evaluierungsprozess durch Mitgestaltungsverfahren zu integrieren.

Bei einer **Empowerment-Evaluation** behandelt der/die Evaluator:in die Stakeholder (z. B. die Mitarbeiter:innen der Initiative, die Teilnehmer:innen) so, als hätten sie die Kontrolle über den Evaluationsprozess, während er/sie die Rolle eines:r Coaches:in oder kritischen Freundes:in übernimmt. Bei dieser Art von Evaluierung sind die Beteiligten letztlich für die wichtigen Entscheidungen (z. B. über die Evaluierungsziele, das Design und die Datenerhebung) zuständig. Sie führen auch die Durchführungs- und Folgemaßnahmen durch und behalten die Kontrolle darüber. Bei einer Empowerment-Evaluation werden die Beteiligten also nicht nur aufgefordert, ihre Bedürfnisse, Werte und Umstände zum Ausdruck zu bringen und zu integrieren, sondern auch die volle Verantwortung für den Evaluationsprozess zu übernehmen.

Zusammengenommen spiegeln kollaborative, partizipative und Empowerment-Evaluierungen die verschiedenen Formen wider, die die Einbeziehung von Stakeholdern in einen Evaluierungsprozess annehmen kann. Während sich die Ansätze vor allem in der Rolle unterscheiden, die sie den Stakeholdern zuweisen, können sie alle genutzt werden, um die Bedürfnisse, Werte und Umstände der Stakeholder in jeder Phase des Evaluierungsprozesses stärker zu berücksichtigen. In dieser Hinsicht verspricht jeder von ihnen, die Inklusivität der Evaluierung zu stärken, Vertrauen zwischen Evaluator:in und Stakeholder:innen zu schaffen und die organisatorische Reflexion und das Lernen zu verbessern. Wenn sie gut umgesetzt werden, können sie alle für die Durchführung einer evidenzbasierten Evaluation geeignet sein.

6. HAUPTARTEN DER EVALUIERUNG



Es gibt unzählige Arten von Evaluierungen, die sich in Bezug auf Ziele, Methoden, Zeitplan und Umfang voneinander unterscheiden. Im Folgenden werden die drei wichtigsten Arten von Evaluierungen vorgestellt. Alle diese Arten können unter Verwendung des im INDEED E-Guidebook 2 beschriebenen evidenzbasierten Evaluierungsansatzes und des INDEED-Evaluierungstools durchgeführt werden.

6.1 FORMATIVE EVALUIERUNG

Die formative Evaluierung wird in der Regel als Teil des Planungsprozesses vor Beginn der Umsetzung durchgeführt, kann aber auch bei laufenden Initiativen durchgeführt werden, wenn diese neu angepasst werden. Sie ist eine Möglichkeit, den Entwurf oder Plan der Initiative zu bewerten.

Die formative Evaluation wirft einen systematischen Blick auf die (geplante) Initiative – was sie bewirken soll, wie ihre Ziele erreicht werden sollen und auf welchen Annahmen sie beruht. Sie kann zeigen, ob der Umsetzungsplan der Initiative die beabsichtigten Ergebnisse erwarten lässt und wie der Plan verbessert werden könnte.

Die formative Evaluierung kann auch in Form einer Pilotimplementierung der geplanten Initiative erfolgen, die es ermöglicht, die programmierten Aufgaben und Aktivitäten zu testen und ihre Korrektheit und Wirksamkeit zu überprüfen.

Was die Methoden und Daten betrifft, so können wissenschaftliche Veröffentlichungen über Radikalisierung und De-Radikalisierung herangezogen werden, um zu bewerten, ob die Annahmen der Initiative durch vorhandene Erkenntnisse gestützt werden. Eine weitere gängige Methode ist die Durchführung von Interviews und Erhebungen unter den wichtigsten Akteur:innen, um ihre Ansichten zum Plan der Initiative zu sammeln. Sie kann auch das Sammeln und Analysieren von Daten über den Kontext der Initiative beinhalten, um festzustellen, ob sie den Bedürfnissen entspricht, die sie erfüllen soll.

Typische Fragen, die formative Evaluierungen beantworten können, sind:

- Stützt sich der Initiativplan auf fundierte Annahmen über Radikalisierung und deren Verhinderung?
- Welche Arten von Aktivitäten sollten durchgeführt werden?
- Entspricht der Plan der Initiative den Bedürfnissen, die mit der Initiative erfüllt werden sollen?
- Funktioniert die Pilotimplementierung so, wie sie sollte?



3 BEISPIEL FÜR EINE FORMATIVE EVALUIERUNG DES PROGRAMMS “VERSTEHEN=RESPEKT”

“UNDERSTAND=RESPECT” ist ein Bildungsprogramm zur Verhinderung von Radikalisierung, die zu Diskriminierung und Hassreden führt, entwickelt von **Dr. Marzena Kordaczuk-Wąs** und umgesetzt von der **Polnischen Plattform für Innere Sicherheit**.

Zielsetzung und Zielgruppe: Ziel der Initiative ist die Förderung eines Verhaltens, das frei von allen Formen von Radikalismus, Diskriminierung und Hassreden ist, bei Schüler:innen und Lehrer:innen in weiterführenden Schulen sowie bei Eltern und in der lokalen Gemeinschaft.

Aktivitäten: Pädagogische Aktivitäten (Workshops, Diskussionen, Thementage), Training von Lebenskompetenzen (psychologisch und sozial, Umgang mit Stress und Aggression, Aufbau von Selbstkontrolle) und Freizeitaktivitäten.

Evaluator:in: Externe:r Evaluator:in (akademische:r Forscher:in)

Ziel der Evaluierung: Die formative Evaluierung wurde während und nach der Pilotphase der Initiative durchgeführt. Ziel war es, herauszufinden, ob die Initiative ordnungsgemäß und effektiv umgesetzt wurde. Die Evaluierung befasste sich auch mit den zugrunde liegenden Annahmen und dem Umsetzungsplan.

Methoden: Interviews mit dem Implementierungsteam des Programms und Schulvertreter:innen

Ergebnisse: Die Initiative wurde allgemein als gut geplant und realistisch beurteilt. Die Evaluierung ergab einige Empfehlungen, wie Inhalt und Organisation weiter verbessert werden könnten.

Für weitere Informationen: Polnische Plattform für Innere Sicherheit. Verstehen = Respektieren (auf Polnisch und Englisch)



6.2 PROZESSEVALUIERUNG

Bei der Prozessevaluierung wird untersucht, wie die Initiative in der Praxis funktioniert. Sie wird in der Regel durchgeführt, um festzustellen, ob die Initiative gemäß dem ursprünglichen Plan umgesetzt wird und um zu erfahren, wie sie verbessert werden kann. Sie kann sich auf verschiedene Aspekte der Umsetzung konzentrieren und viele nützliche Informationen liefern, die zur Verbesserung der Initiative in der Zukunft beitragen können.

Zur Durchführung einer Prozessevaluierung sind Daten über den Durchführungsplan und die tatsächliche Durchführung der Initiative erforderlich. Der Umsetzungsplan der Initiative bestimmt zusammen mit den Gesprächen mit den wichtigsten Interessengruppen die Kriterien, anhand derer die Umsetzung beurteilt werden sollte. Zu den Kriterien kann beispielsweise die Anzahl der organisierten Schulungen oder Veranstaltungen oder die Teilnahmequote an den Schulungen oder Veranstaltungen gehören. Die Kriterien müssen nicht in Zahlen messbar sein, sondern können beispielsweise auch die Zufriedenheit der Teilnehmer:innen mit dem Programm und ihre Ansichten darüber oder die Zusammenarbeit zwischen den verschiedenen an der Durchführung der Initiative beteiligten Partner:innen umfassen.

Typische Fragen, die Prozessevaluierung beantworten können, sind:

- Wie läuft die Umsetzung der Initiative?
- Wird die Initiative (oder ein Teil ihrer Aktivitäten) planmäßig umgesetzt?
- Wie gut hat die Initiative ihre Zielgruppe erreicht?
- Wie erleben die Teilnehmer:innen die Initiative?
- Wie arbeiten die wichtigsten Akteure zusammen, um die Ziele zu erreichen?
- Wie stark sind die wichtigsten Interessengruppen in die Initiative eingebunden?
- Gab es bei der Umsetzung der Initiative irgendwelche Hindernisse?



4 PROZESSEVALUIERUNG DER ARBEIT MEHRERER AGENTUREN (MAW) IN BELGIEN, DEN NIEDERLANDEN UND DEUTSCHLAND

Initiative: Multi Agency Working (MAW) in Belgien, den Niederlanden und Deutschland. Diese behördenübergreifenden Plattformen wurden eingerichtet, um Radikalisierung und gewalttätigen Extremismus zu verhindern.

Evaluator:in: Externe:r Evaluator:in (akademische Forscher:innen am IRCP, Universität Gent)

Ziel der Evaluierung: Die Evaluierung erfolgte in Form einer (realistischen) Prozessevaluierung mit dem Ziel, herauszufinden, "was unter welchen Bedingungen funktioniert". Sie analysierte die Umsetzung des behördenübergreifenden Ansatzes und die Art und Weise der Zusammenarbeit zwischen den



verschiedenen beteiligten Stellen. Dies geschah, um Bereiche mit Verbesserungsbedarf zu ermitteln und zu verstehen, wie der Mechanismus funktioniert.

Methoden der Datenerhebung:

Systematische Literaturrecherche – zur Festlegung von Indikatoren für die Evaluierung.

Feldforschung in einer Stadt pro Land, einschließlich teilnehmende Beobachtung bei MAW-Sitzungen, halbstrukturierte Interviews mit verschiedenen Teilnehmer:innen und Fokusgruppen zur Untersuchung fehlender Elemente.

Methoden der Datenanalyse: Die aus halbstrukturierten Interviews gewonnenen qualitativen Daten wurden anhand eines SWOT-Rahmens (Stärken, Schwächen, Chancen und Risiken) analysiert, um die internen Stärken und Schwächen sowie die externen Chancen und Risiken des MAW zu ermitteln. Die qualitativen Daten aus den Beobachtungen und den Interviews wurden mit den Prozessindikatoren verglichen, die im Rahmen der systematischen Literaturrecherche entwickelt wurden.

Ergebnisse: Der wichtigste Faktor für einen guten behördenübergreifenden Kooperationsprozess war Vertrauen. In allen untersuchten Ländern gab es Probleme mit dem Berufsgeheimnis und dem Untersuchungsgeheimnis sowie einen Mangel an personellen Ressourcen. Die Rolle des Koordinators der MAW-Struktur erwies sich als sehr wichtig. Die Pandemie, der gesellschaftliche Wandel und neue Formen der Radikalisierung wurden als externe Bedrohung angesehen. Als Teil der Ergebnisse und Empfehlungen der Evaluierung wurde ein Online-Selbstevaluierungsinstrument für lokale Praktiker:innen entwickelt (<https://emmascan.eu/>).

Für weitere Informationen: Hardyns, W., Klima, N., & Pauwels, L. (eds.). (2022). Evaluierung und Betreuung des behördenübergreifenden Ansatzes zur Bekämpfung der Radikalisierung in Belgien, den Niederlanden und Deutschland. Antwerpen: Maklu .

6.3 ERGEBNISEVALUIERUNG

Die Ergebnisevaluierung misst die **Auswirkungen** der Initiative. Sie ist eine gängige Methode, um festzustellen, ob die Initiative ihre Ziele erreicht und ein beabsichtigtes Ergebnis erzielt hat. Wie bei der Prozessevaluation müssen Kriterien festgelegt werden, um zu messen, ob die Initiative wirksam war oder nicht. Oft ist es auch notwendig, Informationen über die Situation zu haben, bevor die Umsetzung der Initiative beginnt.

Die Ergebnisevaluierung kann je nach Zielsetzung auf verschiedene Weise durchgeführt werden. Wenn es bei der Evaluierung darum geht, wie die Teilnehmer:innen das Ergebnis der Initiative erleben oder wie viel Veränderung eine Initiative bewirkt hat, sind qualitative Methoden (z. B. Interviews, Umfragen) oft die beste Wahl. Wenn das Ziel darin besteht, nachzuweisen, dass das angestrebte Ergebnis auf die Initiative zurückzuführen ist, können quantitative Methoden verwendet werden, um zu untersuchen, welche Art von Auswirkungen und wie groß diese sind und welches Merkmal oder welche Dimension der Intervention am vorteilhaftesten zu sein scheint. Je nach verwendeter Methode variiert der Grad der Evidenz der quantitativen Ergebnisevaluierung (d. h. wie zuverlässig die Schlussfolgerungen sind, die auf dieser Grundlage gezogen werden können). Die stärkste Evidenz wird mit randomisierten kontrollierten Studien erzielt.

Es ist wichtig, darauf hinzuweisen, dass quantitative und qualitative Evaluierungsmethoden sich gegenseitig ergänzen. Allerdings kann die eine die andere nicht ersetzen. Qualitative Methoden können keine kausalen Schlussfolgerungen über

die Wirkung der Intervention ziehen. Andererseits ist es mit quantitativen Methoden im Allgemeinen nicht möglich zu untersuchen, wie Einzelpersonen die Intervention erlebt haben.

Typische Fragen, die durch Ergebnisevaluation beantwortet werden können, sind:

- Welche Auswirkungen hat die Initiative?
- Hat die Initiative ihre Ziele und Ergebnisse erreicht?
- Inwieweit hat sich das Wissen, die Fähigkeiten, die Einstellungen und das Verhalten der Zielgruppe nach der Teilnahme an der Initiative verändert? Wie verhält sich dies im Vergleich zu den Veränderungen, die bei denjenigen beobachtet wurden, die nicht an der Initiative teilgenommen haben?
- Gab es unbeabsichtigte Auswirkungen auf die Zielgruppe oder den Kontext?



5 ERGEBNISEVALUIERUNG VON AGGREDI IN HELSINKI, FINNLAND

Initiative: Aggredi

Ziele der Initiative und Zielgruppe: Aggredi ist eine Initiative einer NRO (HelsinkiMissio), die darauf abzielt, die Rückfälligkeit von 18- bis 49-jährigen Straftätern zu verringern, die wegen Straßengewalt verurteilt wurden. Sie bietet Mentoring und praktische Hilfe mit dem Ziel, die Wiedereingliederung in die Gesellschaft zu unterstützen.

Evaluator:in: Forscher:innen des Nationalen Forschungsinstituts für Rechtspolitik in Finnland.

Ziel der Evaluierung: Herausfinden, ob es der Initiative gelungen ist, die Rückfälligkeit ihrer Klienten zu verringern.

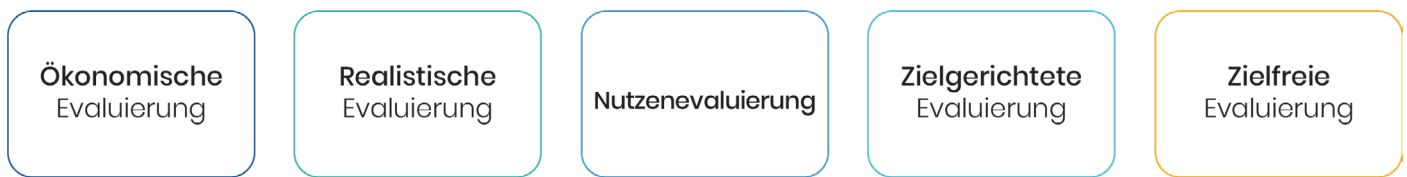
Methode: Eine quantitative Studie, in der die Rückfälligkeit verglichen wurde zwischen 1) Klienten, die das Aggredi-Programm abgeschlossen haben, 2) Klienten, die daran teilgenommen, es aber abgebrochen haben, 3) Straftätern, denen die Teilnahme angeboten wurde, die sie aber abgelehnt haben, und 4) einer Vergleichsgruppe ähnlicher Straftäter, die keinen Kontakt mit Aggredi hatten.

Ergebnisse: Die Evaluierung ergab, dass die Rückfallquote bei denjenigen, die das Programm abgeschlossen haben, deutlich geringer ist als bei denjenigen, die es abgebrochen haben. Außerdem zeigte sich, dass die Rückfallquote umso geringer war, je länger man an dem Programm teilnahm. Aufgrund methodischer Schwierigkeiten war es schwierig, schlüssig nachzuweisen, dass die beobachtete geringere Rückfallquote auf die Programmteilnahme zurückzuführen ist.

Für weitere Informationen: [Evaluationsbericht des Aggredi-Programms](#) (auf Finnisch).



6.4 ANDERE EVALUIERUNGSARTEN



Neben den drei oben genannten allgemeinen Evaluationsarten werden oft noch viele weitere spezifische Evaluationsarten genannt. Im Folgenden sind einige Beispiele aufgeführt:

Wirtschaftliche Evaluierung

Die wirtschaftliche Evaluierung ist nützlich, um das Verhältnis zwischen den **Kosten** und der **Wirksamkeit** der Initiative zu untersuchen. So kann beispielsweise eine sehr wirksame Initiative sehr kostspielig und daher nicht umsetzbar sein, oder eine Initiative kann geringe Kosten verursachen, aber unwirksam sein.

Diese Form der Evaluierung wird meist von Geldgeber:innen genutzt, um den finanziellen Wert der Initiative zu ermitteln, und von politischen Entscheidungsträger:innen, um die Kosten verschiedener Initiativen zu ermitteln und zu vergleichen. Außerdem können die Verwalter:innen und Durchführenden der Initiative die Intervention ändern, wenn sie zu viele Kosten verursacht.

Typische Fragen, die wirtschaftliche Evaluierungen beantworten können, sind:

- Welche Mittel werden für die Initiative eingesetzt?
- Wie werden die Ressourcen genutzt?
- Wie werden die Kosten in Ergebnisse umgewandelt?

Realistische Evaluierung

Die realistische Evaluierung ist eine Art von Evaluierung, die sich auf das "Wie" und "Warum" der Ergebnisse einer Initiative konzentriert, indem sie die zugrunde liegenden Mechanismen ermittelt, die dazu führen, dass Initiativen bestimmte Ergebnisse erzielen. Sie ist besonders nützlich, um zu verstehen, wie eine Initiative in verschiedenen Kontexten und unter welchen Bedingungen funktioniert, da sie zu erklären versucht, warum eine Initiative für einige funktioniert, für andere aber nicht.

Typische Fragen, die realistische Evaluierungen beantworten können, sind:

- Welches sind die Kontextfaktoren, die die Wirksamkeit der Initiative beeinflussen?
- Welche Initiativstrategien sind am wirksamsten?
- Welche Aktivitäten haben die größte Wirkung?
- Welche Faktoren machen die Initiative in bestimmten Kontexten erfolgreich und in anderen nicht?

Nutzenevaluierung

Die nutzungsorientierte Evaluation basiert auf der Idee, dass die Evaluationsergebnisse für die primär vorgesehenen Nutzer:innen nützlich sein müssen und die Ergebnisse genutzt werden sollten. Aus diesem Grund sollte diese Art der Evaluierung so konzipiert sein, dass die Ergebnisse genutzt werden, und die Nutzer der Evaluierung müssen bereits in der Planungsphase der Evaluierung einbezogen werden.

Die Nutzungsevaluierung kann in Kombination mit den zuvor beschriebenen formativen, Prozess- und Ergebnisevaluierungen und anderen Methoden verwendet werden, je nach den Bedürfnissen der vorgesehenen Nutzer:innen. Die Methode hängt von den Daten ab, die für die Beantwortung der zentralen Evaluierungsfragen benötigt werden, sowie von der geeignetsten Methode, um diese Ergebnisse zu liefern.

Zielgerichtete Evaluierung

Die zielorientierte Evaluierung misst, ob die Initiative Fortschritte **in Richtung einer Reihe von Zielen macht**, die SMART (spezifisch, messbar, erreichbar, relevant und zeitgerecht) sind. Diese Evaluierung dient den Geldgeber:innen oder Durchführenden der Initiative als Nachweis dafür, dass sich die Initiative in Richtung der Ziele bewegt, die in der Planungsphase der Initiative vereinbart wurden. Die gesetzten Ziele werden nicht in Frage gestellt.

Typische Fragen, die durch zielgerichtete Evaluierungen beantwortet werden können, sind:

- Hat die Initiative ihre Vorgaben und Ziele erreicht?
- Wurden die Ziele als Ergebnis der Initiative oder aufgrund anderer externer Faktoren erreicht?

Zielfreie Evaluierung

Bei der zielfreien Evaluierung werden die Ergebnisse der Initiative untersucht, ohne sich auf ein einzelnes Ziel zu konzentrieren. Der/die Evaluator:in kennt nicht einmal die Ziele der Initiative. Die Evaluierung analysiert die Ergebnisse und Auswirkungen der Initiative, ohne sich von einem bestimmten, vorher festgelegten Schwerpunkt leiten zu lassen. Auf diese Weise ist der/die Evaluator:in weder voreingenommen noch von irgendwelchen Erwartungen beeinflusst. Der/die Evaluator:in muss ein:e Externe:r sein und darf weder die Organisation noch die Initiative kennen.

Die zielfreie Evaluierung untersucht die unbeabsichtigten Ergebnisse und Veränderungen, die durch die Initiative verursacht werden, und kann die positiven und negativen Nebeneffekte ermitteln. Sie wird eingesetzt, wenn die Initiative die Arbeit in einem komplexen Umfeld erfordert und die Ziele unklar sind.

Typische Fragen, die eine zielfreie Evaluation beantworten kann, sind:

- Welche Ergebnisse hat die Initiative erbracht?
- Können die Ergebnisse auf die Initiative zurückgeführt werden?
- Sind die Auswirkungen positiv oder negativ?

6.5 KOMBINATION VON EVALUIERUNGSARTEN

Es ist möglich, verschiedene Arten von Evaluierungen in einer einzigen Evaluierung zu kombinieren. Dies erfordert mehr Ressourcen, kann aber auch nützlichere und umfassendere Ergebnisse liefern. Die Kombination von Ergebnis- und Prozessevaluierung ist durchaus üblich und kann vor allem dann hilfreich sein, wenn man eine Erklärung dafür sucht, warum die Initiative die gewünschte Wirkung erzielt hat oder nicht. Eine Kombination aus Prozess- und Ergebnisevaluierung kann z. B. dabei helfen zu ermitteln, inwieweit das Nichterreichen des gewünschten Ergebnisses auf eine mangelhafte Umsetzung oder Schwächen in der Konzeption zurückzuführen ist.

Mehr erfahren

Für weitere Informationen über formative, Prozess- und ErgebnisEvaluierungen:

- Werkzeugkasten für die Evaluierung. [Arten der Evaluierung](#).
- BessereEvaluierung. [Was ist Evaluation?](#)
- James Bell Associates. (2018). [Formative Evaluation Toolkit: Ein Schritt-für-Schritt-Leitfaden und Ressourcen für die Evaluierung der Programmumsetzung und der ersten Ergebnisse](#). Washington, DC: Children's Bureau, Administration for Children and Families, U.S. Department of Health and Human Services.
- Thompson, S. & Leroux, E. (2023). Lessons learned from dual site formative evaluations of Countering violent extremism (CVE) programming coiled by Canadian police. *Journal of Policing, Intelligence and Counter Terrorism*, 18(1).

Für weitere Informationen über andere Arten von Evaluierungen:

- MEASURE Evaluierung. [Wirtschaftliche Evaluierung](#).
- WHO. [Ökonomische Evaluierungen](#).
- BessereEvaluierung. Realistische Evaluierung.
- BetterEvaluation. [Verwendungsorientierte Evaluierung](#).
- Patton, M. Q. (2008). *Nutzungsorientierte Evaluation* (4. Auflage). Thousand Oaks, CA: Sage.
- Youker, B. W., & Ingraham, A. (2014). Goal-free evaluation: Eine Orientierung für die Evaluierung von Stiftungen. *The Foundation Review*, 5(4).



7. EVALUIERUNGSKONZEPTE UND –METHODEN

Bei einer Evaluierung müssen zahlreiche Entscheidungen und Auswahlmöglichkeiten getroffen werden. Eine wichtige Entscheidung bei der Durchführung einer Evaluierung ist die Wahl des geeigneten Evaluierungsdesigns und der Methoden. Die Art der Evaluierung gibt die Richtung vor, legt aber noch nicht fest, wie die Evaluierung durchgeführt werden soll. Das Evaluationsdesign legt die Gesamtstruktur und den Umfang der Evaluierung fest. Die Evaluierungsmethoden ergänzen es, indem sie festlegen, wie die Daten gesammelt und analysiert werden sollen (siehe Tabelle 1).

Tabela 1: Kluczowe decyzje w planowaniu ewaluacji

Art der Evaluierung	
Ausgewählt auf der Grundlage der Ziele der Evaluierung Formativ, Prozess, Ergebnis, etc.	
Entwurf der Evaluierung	
Legt den Umfang und die Struktur der Evaluierung fest	
Methoden der Datenerhebung	Methoden der Datenanalyse
Welche Art von Daten wird verwendet und wie werden sie erhoben?	Wie die gesammelten Daten analysiert werden, um Schlussfolgerungen zu ziehen

Welches Evaluierungsdesign und welche Methoden am besten geeignet sind, hängt von den Zielen, dem Umfang und den für die Evaluierung verfügbaren Ressourcen ab. Dieses Kapitel gibt einen kurzen Überblick über einige gängige Evaluationsdesigns und Methoden zur Datenerhebung und -analyse.

7.1 EVALUIERUNGSENTWÜRFE

Das Evaluationsdesign legt die **Grundstruktur der Evaluation** fest. Es legt zum Beispiel fest, ob sich die Evaluation auf die Analyse der Situation zu einem bestimmten Zeitpunkt konzentriert oder einen Vergleich über verschiedene Zeiträume beinhaltet. Eine weitere wichtige Entscheidung ist, ob sich die Evaluierung nur auf die Initiative selbst konzentriert oder ob auch Vergleichsgruppen herangezogen werden.

Einige Evaluationsdesigns sind recht einfach umzusetzen, während andere mehr Fachwissen und tiefgreifende Kenntnisse wissenschaftlicher Methoden erfordern. In diesem Abschnitt werden einige der am häufigsten verwendeten Evaluationsdesigns vorgestellt, von weniger anspruchsvollen bis hin zu anspruchsvolleren.

Gestaltung der Fallstudie

Bei einer Evaluierung in Form einer Fallstudie werden **detaillierte Informationen über eine kleine Anzahl von Fällen** (z. B. einige Teilnehmer:innen einer Initiative) gesammelt. Sie können detaillierte und reichhaltige Erkenntnisse über die Funktionsweise und (unbeabsichtigten) Folgen einer Initiative liefern. Es ist wichtig zu erkennen, dass Fallstudien auch einige bemerkenswerte Einschränkungen haben. Die **aus einer Fallstudie gewonnenen Erkenntnisse sind in der Regel nicht repräsentativ oder verallgemeinerbar**. Es ist selten möglich zu wissen, ob die Erfahrungen einer kleinen Anzahl von Teilnehmer:innen, die für die Evaluierung befragt wurden, die Erfahrungen der Teilnehmer:innen von Initiativen im Allgemeinen widerspiegeln.

Die Fallstudie ist **besonders nützlich, wenn der Zugang zu Daten begrenzt ist**. Manchmal können Fallstudien das einzige verfügbare Evaluierungsdesign sein. Fallstudien sind auch sehr nützlich für die Formulierung von Hypothesen und Evaluierungsfragen, die später mit komplexeren Evaluierungsdesigns untersucht werden können. Da sie tiefere Einblicke in die Funktionsweise der Initiative bieten, können sie auch eine gute Ergänzung zu Evaluierungen sein, die ansonsten quantitative Methoden und numerische Indikatoren verwenden.



6 BEISPIEL FÜR EIN FALLSTUDIENDESIGN

Cherney und Belton (2020) entschieden sich für ein Fallstudiendesign, um das Programm zur sozialen Wiedereingliederung (PRISM) zu untersuchen. Sie konzentrierten sich auf nur drei Klienten, für die ausreichend Daten zur Verfügung standen, um die Entwicklung im Laufe der Zeit zu untersuchen.

Bei der Evaluierung wurden die Auswirkungen von PRISM durch die Erhebung von Informationen über die Ansichten der Betroffenen zu verschiedenen Zeitpunkten bewertet. Die erforderlichen Informationen wurden durch verschiedene Datenerhebungsmethoden gesammelt, darunter Interviews mit den PRISM-Mitarbeiter:innen und den Einzelpersonen selbst, die Überprüfung der von den PRISM-Psycholog:innen ausgefüllten Fortschrittsberichte und die von den PRISM-Mitarbeiter:innen erstellten Fallnotizen der Kunden. Zur Analyse der Daten wurde eine qualitative Inhaltsanalyse durchgeführt. Die Ergebnisse zeigten, wo die einzelnen Klienten Fortschritte gemacht hatten und wo noch Verbesserungen erforderlich waren.

Für weitere Informationen: Cherney, A., & Belton, E. (2020). Evaluierung von Interventionsergebnissen bei radikalisierten Straftätern: Testing the pro integration model of extremist disengagement as an evaluation tool. *Dynamics of Asymmetric Conflict*, 13(3), 193-211.

Querschnittstypen

Bei einer Evaluierung mit einem Querschnittsdesign werden Daten von möglichst vielen Personen (z. B. Personen, die eine Initiative durchführen oder an ihr teilnehmen) zu einem bestimmten Zeitpunkt und in der Regel mithilfe einer Umfrage erhoben. Die Verwendung eines solchen Designs ermöglicht die Erhebung von Daten aus einem großen Pool von Personen und den Vergleich von Unterschieden zwischen Gruppen. So kann beispielsweise analysiert werden, ob sich die Personen, die an der Initiative teilgenommen haben, von den anderen Studienteilnehmer:innen in Bezug auf bestimmte Aspekte im Zusammenhang mit den Zielen der Initiative (z. B. Einstellungen oder Verhalten) unterscheiden.

Daher kann ein Querschnittsdesign **eine nützliche “Momentaufnahme” der Funktionsweise, der Rezeption oder der potenziellen Auswirkungen der**



Initiative darstellen. Es eignet sich besonders für Prozessevaluierungen. Für Ergebnisevaluierungen ist es nicht die ideale Option. Dies liegt daran, dass es aufgrund des kurzen Zeitrahmens schwierig ist, die Auswirkungen zu bewerten, und dass das Fehlen einer Kontrollgruppe die Fähigkeit einer Evaluierung einschränkt, einen kausalen Zusammenhang zwischen einer Initiative und ihren Auswirkungen herzustellen. Darüber hinaus können die Unterschiede zwischen den Teilnehmer:innen der Initiative und den anderen Studienteilnehmer:innen auch auf andere Faktoren als die Initiative und ihre Auswirkungen zurückzuführen sein (d. h. Selektionsverzerrungen). Die Personen, die an der Initiative teilnehmen, unterscheiden sich wahrscheinlich in vielerlei Hinsicht von anderen, was in der Forschung möglicherweise nicht berücksichtigt werden kann.



7 BEISPIEL FÜR EIN QUERSCHNITTMUSTER

Dunn et al. Bewerteten die Strategie der australischen NSW-Polizei zur Bekämpfung von Radikalisierung, die eine Initiative zur Einbindung muslimischer Gemeinschaften in Sydney umfasste. Die Evaluator:innen waren akademische Forscher:innen, die mit internen Mitarbeiter:innen der NSW-Polizei zusammenarbeiteten. Die Initiative wurde durch die Analyse der Daten einer einmaligen Umfrage bewertet, die unter den Mitgliedern der von der Initiative betroffenen Gemeinschaften durchgeführt wurde. Im Rahmen der Umfrage wurden Daten über das Bewusstsein der Muslime in Sydney für die Initiative der NSW-Polizei zur Einbindung in die Gemeinschaft und über ihre Erfahrungen mit der Initiative sowie über das Vertrauen, die Zusammenarbeit und die Wahrnehmung der NSW-Polizei erhoben.

Die Studie hat gezeigt, dass es der Initiative gelungen ist, direkte Kontakte mit der Gemeinschaft, dauerhafte Beziehungen und ein starkes Gemeinschaftsbewusstsein aufzubauen. Die Befragten empfahlen, die Kontakte zu intensivieren, die Sichtbarkeit zu verbessern und die Partnerschaften zu stärken.

Für weitere Informationen: Dunn, K. M., Atie, R., Kennedy, M., Ali, J. A., O'Reilly, J., & Rogerson, L. (2016). Can you use community policing for counter terrorism? Evidence from NSW, Australia. *Police Practice and Research*, 17(3), 196–211.

Längsschnittsdesign

Das Längsschnittsdesign wird in der Regel verwendet, um **die Auswirkungen einer Initiative im Lauf der Zeit zu bewerten**. Bei der häufigsten Form einer Längsschnittstudie erhebt der/Die Evaluator:in Daten von Teilnehmer:innen zu zwei oder mehreren Zeitpunkten vor und nach dem Beginn einer Initiative. Die Anwendung eines solchen Längsschnittsdesigns kann ein leistungsfähiges Instrument zur Evaluierung der kurz- und langfristigen Auswirkungen einer Initiative sowie ihres ordnungsgemäßen Funktionierens sein. Da die Daten zu verschiedenen Zeitpunkten erhoben werden müssen, ist für die Durchführung ein längerer Zeitrahmen erforderlich. Auf der Grundlage eines solchen Längsschnittsdesigns können jedoch keine kausalen Schlussfolgerungen über die Wirksamkeit der Initiative gezogen werden. Die wichtigste Einschränkung betrifft das Fehlen einer kontrafaktischen Dimension (die Unmöglichkeit zu wissen, was passiert wäre, wenn die Person nicht an der Initiative teilgenommen hätte).



TIPP:

Im Idealfall wird die Evaluierung mit einem Längsschnittsdesign zusammen mit der Initiative selbst geplant. Dadurch wird sichergestellt, dass die erforderlichen Daten zum richtigen Zeitpunkt erhoben werden und die Überwachungspraktiken der Initiative die Evaluierung unterstützen.





8 BEISPIEL FÜR EIN LÄNGSSCHNITTDESIGN

Die Wissenschaftler Feddes, Mann und Doosje haben ein niederländisches Resilienztrainingsprogramm in einem Längsschnittdesign ausgewertet. Insgesamt 46 junge Muslime mit Migrationshintergrund nahmen an dem Training teil. Für die Evaluierung wurden Daten von den Programmteilnehmern mittels eines Fragebogens zu vier Zeitpunkten erhoben: 1) vor Beginn des Trainings, 2) nach Abschluss des ersten Moduls, 3) unmittelbar nach Abschluss des Trainings und 4) drei Monate nach dem Training.

Die Ergebnisse der Evaluierung zeigen, dass sich das Training positiv auf die Teilnehmer auswirkte und ihr Selbstwertgefühl, ihr Einfühlungsvermögen und ihre Fähigkeit, das Verhalten und die Reaktionen anderer Menschen vorherzusehen, verbesserte. Außerdem berichteten die Teilnehmer über geringere Gewaltabsichten und eine geringere Einstellung zu ideologiebasierter Gewalt. Die Ergebnisse zeigten jedoch, dass die Teilnehmer ein höheres Maß an Narzissmus aufwiesen. Insgesamt zeigten die Ergebnisse, dass die Stärkung der Teilnehmer und insbesondere die Förderung ihrer Empathie zu einem erfolgreichen Resilienztraining führen.

Für weitere Informationen: Feddes, A. R., Mann, L., & Doosje, B. (2015). Steigerung des Selbstwertgefühls und der Empathie zur Verhinderung gewalttätiger Radikalisierung: eine quantitative Längsschnittauswertung eines Resilienztrainings, das sich auf Jugendliche mit einer doppelten Identität konzentriert. Zeitschrift für angewandte Sozialpsychologie, 45(7), 400–411.

(Quasi-)experimentelle Designs

Bei einer Evaluierung mit einem (quasi-)experimentellen Design wird **eine Gruppe von Personen (oder Schulen, Gefängnissen, Städten usw.), auf die eine Initiative angewendet wurde, mit einer Kontrollgruppe verglichen, in der diese Initiative nicht angewendet wurde.** Die Verwendung einer Kontrollgruppe ermöglicht es, zu analysieren, ob die beobachtete Veränderung bei den Teilnehmer:innen wirklich auf die Initiative und nicht auf andere Faktoren zurückzuführen ist.

Eine Kontrollgruppe kann auf unterschiedliche Weise gebildet werden. In einer experimentellen Studie (auch **randomisierte kontrollierte Studie** oder RCT genannt) werden die Teilnehmer:innen nach dem Zufallsprinzip einer Behandlungs- und einer Kontrollgruppe zugeteilt. Im Bereich der PVE/CVE ist eine solche zufällige Zuteilung sowohl aus praktischen als auch aus ethischen Gründen schwierig. Unter dem Gesichtspunkt des Kausalitätsschlusses liefern randomisierte Experimente den stärksten Beweis für die Wirksamkeit der Initiative.



9 BEISPIEL FÜR EINE RANDOMISIERTE KONTROLLIERTE STUDIE

Bei der Evaluierung eines dänischen Regierungsprojekts zur Bekämpfung extremistischer Narrative unter Jugendlichen wurde eine randomisierte kontrollierte Studie durchgeführt. Die Evaluierung wurde von den Wissenschaftlern Parker und Lindekilde durchgeführt. Sie erstellten ein Umfrageexperiment mit einer Kontroll- und einer Behandlungsgruppe. Die Kontrollgruppe hatte keinen Kontakt mit dem Projekt, die Behandlungsgruppe hatte Kontakt mit dem Projekt. Die Teilnehmer:innen der Kontrollgruppe beantworteten eine Umfrage, bevor sie an einem Workshop mit ehemaligen Extremisten teilnahmen, während die Teilnehmer:innen der Behandlungsgruppe die Fragen direkt nach dem Besuch desselben Workshops beantworteten. Um zu bewerten, wie wirksam die Initiative im Hinblick auf die Verringerung der Unterstützung politischer Gewalt war, stellten die Evaluator:innen indirekte Fragen zu politischer Gewalt.

Für weitere Informationen: Parker, D., & Lindekilde, L. (2020). Extremismusprävention mit Extremisten: Ein zweischneidiges Schwert? Eine Analyse der Auswirkungen des Einsatzes ehemaliger Extremisten in dänischen Schulen. Erziehungswissenschaften, 10(4).



Eine weitere Alternative ist ein **quasi-experimentelles Design**, bei dem die Kontrollgruppe auf andere Weise als durch Zufallszuweisung gebildet wird. So kann ein:e Evaluator:in beispielsweise eine Kontrollgruppe bilden, indem er eine Population in einem bestimmten Umfeld (z. B. Schule, Stadt, Gefängnis), in dem eine Initiative durchgeführt wurde, mit einer Population in einem vergleichbaren Umfeld vergleicht, in dem die Initiative nicht durchgeführt wurde (z. B. eine nahe gelegene Schule, Stadt oder ein Gefängnis mit ähnlichen Merkmalen).



10 BEISPIEL FÜR EIN QUASI-EXPERIMENTELLES DESIGN

Die Evaluierung des Aggredi-Programms wurde durchgeführt, um herauszufinden, ob es der Initiative gelungen war, die Rückfälligkeit ihrer Klient:innen zu verringern. Das Hauptziel der Verwendung eines quasi-experimentellen Designs bestand darin, festzustellen, ob es der Initiative gelungen war, die Rückfälligkeit zu verringern, indem abgeschätzt wurde, was mit den Personen, die an der Behandlung teilgenommen hatten, geschehen wäre, wenn sie nicht am Aggredi-Programm teilgenommen hätten.

Zu diesem Zweck wurde eine quantitative Studie durchgeführt, in der die Rückfälligkeit zwischen drei Behandlungsgruppen und einer Kontrollgruppe verglichen wurde. Die Behandlungsgruppen waren: 1) Klient:innen, die das Aggredi-Programm abgeschlossen haben, 2) Klient:innen, die teilgenommen, aber abgebrochen haben, und 3) Straftäter:innen, denen die Teilnahme angeboten wurde. Für alle drei Gruppen gab es eine Vergleichsgruppe von ähnlichen Straftäter:innen, die keinen Kontakt zu Aggredi hatten.

Die Grundidee bestand darin, eine Vergleichsgruppe zu bilden, um die Rückfälligkeit der Behandlungsgruppe mit der Rückfälligkeit der Vergleichsgruppe zu vergleichen, die aus Personen bestand, die den Aggredi-Teilnehmer:innen in Bezug auf ihr Alter, ihr Geschlecht und ihren kriminellen Hintergrund ähnlich waren. Die Vergleichsgruppe wurde auf der Grundlage von Daten aus dem Forschungsregister für Straftaten und Sanktionen des Forschungsinstituts für Rechtspolitik gebildet.

Für weitere Informationen: [Evaluierungsbericht des Aggredi-Programms](#) (auf Finnisch).

7.2 METHODEN DER DATENERHEBUNG

Evaluierungskonzepte schreiben in der Regel keine bestimmten Datenerhebungsmethoden vor, und sie können mit einer Reihe von Methoden kombiniert werden. In diesem Abschnitt werden einige wichtige Datenerhebungsmethoden vorgestellt, die bei Evaluierungen verwendet werden können.

Systematische Überprüfung der vorhandenen Dokumentation

Ein guter Ausgangspunkt für die Datenerhebung ist die Prüfung, **welche Art von Daten bereits über die Initiative verfügbar sind**. Zuden relevanten vorhandenen Datengehören die Umsetzungspläne der Initiative, Fortschrittsberichte und alle Dokumente, die die Ziele und die Theory of Change erläutern. Dazu können auch Daten zur Projektüberwachung gehören, wie z. B. Aufzeichnungen über Aktivitäten und Teilnehmer:innen.

Vorhandene Dokumente sind hilfreich für das Verständnis der Ziele und Pläne der Initiative. Monitoring-Daten können eine hervorragende Quelle für Längsschnittstudien sein, bei denen es wichtig ist, Daten von verschiedenen Zeitpunkten zu haben, um Veränderungen diachron zu beobachten. Im Idealfall sind die Überwachungspraktiken der Initiative von Anfang an so angelegt, dass sie die für die Evaluierung erforderlichen Daten liefern.

Interviews

Interviews sind eine häufig verwendete Methode zur Datenerhebung bei der Evaluierung von PVE/CVE/De-rad-Initiativen. Sie sind besonders hilfreich, um **eingehende Informationen und Einblicke in die Umsetzung der Initiative zu erhalten**. Sie werden häufig auch zur **Evaluierung spezifischer Auswirkungen** der Initiative verwendet (z. B. Auswirkungen auf die Ansichten und Einstellungen der Teilnehmer).

Interviews können auf unterschiedliche Weise durchgeführt werden. Sie können strukturiert sein, was bedeutet, dass alle Befragten die gleichen Fragen beantworten. Eine andere Alternative ist ein halbstrukturiertes Interview, bei dem die Befragten zu denselben Themen befragt werden, die Reihenfolge der Fragen jedoch geändert werden kann und der/die Interviewer:in Folgefragen stellen kann. Schließlich können die Interviews auch unstrukturiert sein und die Form eines freien Gesprächs annehmen.

Jede Art von Befragung hat ihre eigenen Vor- und Nachteile. Strukturierte Interviews sind am besten geeignet, wenn es darauf ankommt, leicht vergleichbare Antworten zu erhalten. Wenn es darauf ankommt, ein tieferes Verständnis der Wahrnehmungen eines:r Befragten zu erlangen, ist es in der Regel besser, halbstrukturierte oder unstrukturierte Interviews zu verwenden, da sie dem:r Befragten mehr Raum geben, seine/ihre Ansichten in eigenen Worten zu erklären.

Schwerpunktgruppen

Fokusgruppen sind Diskussionen mit einer Gruppe von (in der Regel fünf bis zehn) Personen über ein bestimmtes Thema. Sie können als **Alternative zu Interviews** dienen **oder diese ergänzen**. Die Teilnehmer:innen der Fokusgruppen können eine bestimmte Gruppe von Personen repräsentieren (z. B. Teilnehmer:innen an einer Initiative) oder aus einer Vielzahl von Interessengruppen bestehen (z. B. Praktiker:innen, die eine Initiative auf verschiedenen Ebenen umsetzen).

Fokusgruppen können eine kostengünstige und zeitsparende Alternative zu ausführlichen Einzelinterviews sein. Sie sind auch eine gute Möglichkeit, Gemeinsamkeiten und Unterschiede bei den Erfahrungen mit der Umsetzung und den Auswirkungen einer Initiative herauszuarbeiten.

Erhebungen

Eine Umfrage ist eine weitere Möglichkeit, Informationen von mehreren Befragten zu sammeln. Bei einer Umfrage wird eine Reihe von Personen gebeten, einen standardisierten Fragebogen auszufüllen. Umfragen bieten eine **ressourceneffiziente Möglichkeit, innerhalb einer kurzen Zeitspanne Daten von einer großen Gruppe von Personen zu sammeln**, um beispielsweise die Auswirkungen einer Initiative auf

die Teilnehmer:innen zu bewerten. Der Fragebogen kann eigenständig oder als angeleitete Aktivität in einer persönlichen oder Online-Umgebung ausgefüllt werden.

Erhebungen sind zwar ein effizientes Mittel zur Datenerfassung, haben aber auch erhebliche **Einschränkungen**. Ein sehr häufiges Problem ist die niedrige Antwortquote (nur eine begrenzte Anzahl von Befragten füllt den Fragebogen aus). Außerdem hängt der Nutzen der Umfrage in hohem Maße davon ab, wie gut die Fragen formuliert sind. Es ist wichtig, dass die Fragen eindeutig und für die Befragten leicht zu verstehen sind. Es wird dringend empfohlen, viel in die Entwicklung des Fragebogens zu investieren und ihn zu testen, bevor er zum Einsatz kommt.

Was an Effizienz und Anzahl der Befragten gewonnen wird, geht an Detailgenauigkeit und Tiefe verloren. Mit einem Fragebogen erhobene Daten sind oft recht begrenzt und oberflächlich. Wenn es also darauf ankommt, vertieftes Wissen zu sammeln, sind Interviews oder Fokusgruppen in der Regel die bessere Alternative.

Teilnehmende und nichtteilnehmende Beobachtung

Die teilnehmende Beobachtung bezieht sich auf einen Datenerhebungsprozess, bei dem der/die Evaluator:in direkt mit der Initiative in Kontakt tritt und an deren täglichen Aktivitäten teilnimmt. Nichtteilnehmende Beobachtung bedeutet, dass die Aktivitäten der Initiative beobachtet werden, ohne aktiv daran teilzunehmen.

Die teilnehmende und nichtteilnehmende Beobachtung **ermöglicht es dem/r Evaluator:in, die Umsetzung der Initiative aus erster Hand zu erleben**. Sie ermöglichen ein detailliertes Verständnis der täglichen Abläufe einer Initiative, einschließlich ihrer Herausforderungen und (unbeabsichtigten) Folgen. Sie können besonders nützlich sein, wenn sie in Kombination mit anderen Datenerhebungsmethoden eingesetzt werden, und sie können eine Grundlage für die fundierte Vorbereitung von Tiefeninterviews oder Fragebögen bieten.



II BEISPIEL: DATENERHEBUNGSMETHODEN BEI DER EVALUIERUNG VON STRIVE II (STRENGTHENING RESILIENCE TO VIOLENT EXTREMISM), 2017–2020

Die vom Royal United Services Institute (RUSI) durchgeführte STRIVE II-Initiative zielt darauf ab, die Rekrutierung und Unterstützung gewalttätiger extremistischer Gruppen sowie die Radikalisierung an Brennpunkten in Kenia insgesamt zu verringern. Die Initiative konzentriert sich auf die Bekämpfung struktureller Faktoren, gruppenbezogener Dynamiken und individueller Faktoren, die die Bedingungen für Radikalisierung und Rekrutierung schaffen und zu diesen beitragen.

Die Evaluierung wurde von RUSI in Auftrag gegeben und von externen Evaluator:innen durchgeführt, bei denen es sich um akademische Forscher:innen handelte. Die Evaluierung kombinierte eine Ergebnisevaluierung der Auswirkungen der Initiative bis zu diesem Zeitpunkt mit einer formativen Evaluierung, um die künftigen Maßnahmen zu beurteilen.

Bei der Evaluierung wurden mehrere Methoden der Datenerhebung eingesetzt:
Vorhandene Dokumente – 40 Hintergrunddokumente wie Managementberichte, politische Dokumente, veröffentlichte Forschungs- und Programmleitlinien, Überwachungs- und Evaluierungsdaten, Evaluierungs- und Monitoringleitlinien sowie Dokumente, die die Theory Of Change der Initiative

beschreiben (wie die beabsichtigten Ergebnisse erreicht werden sollten).

Interviews und Fokusgruppen – mehr als 50 halbstrukturierte Interviews und fünf Fokusgruppen mit Mitarbeiter:innen der Initiative, Teilnehmer:innen und Vertreter:innen der Geldgeber:innen und Partner:innen.

Nichtteilnehmende Beobachtung – Beobachtung verschiedener Aktivitäten der Initiative, einschließlich Schulungen, interner Treffen und öffentlicher Veranstaltungen.

Für weitere Informationen: Fisher, T., Range, D., & Cuddihy, J. (2020). [Evaluierung von 'Violent Extremism Strengthening Resilience \(STRIVE II\)' in Kenia: Final report.](#)

(Akademische) Literatur

Eine weitere wichtige Datenquelle für die Evaluierung ist die vorhandene (akademische) Literatur, die fast immer nützlich ist, um Evaluierungen ähnlicher Initiativen zu ermitteln und geeignete Analysemethoden und Indikatoren zu bestimmen. Sie ist unverzichtbar, wenn das Ziel darin besteht, zu bewerten, inwieweit die Theory of Change und die Arbeitsmethoden der Initiative mit den vorhandenen wissenschaftlichen Erkenntnissen übereinstimmen. Dies ist ein übliches Ziel bei formative Evaluierungen.



12 BEISPIEL FÜR DIE VERWENDUNG WISSENSCHAFTLICHER LITERATUR

Der flämische Aktionsplan zur Prävention von Radikalisierung und Polarisierung wurde vom flämischen Friedensinstitut bewertet. Die Evaluierung analysierte den Aktionsplan im Allgemeinen und konzentrierte sich speziell auf mehrere durch den Aktionsplan finanzierte Projekte. Das flämische Friedensinstitut führte eine eingehende Analyse der konkreten Aktionsbereiche des Aktionsplans durch. Um zu bewerten, ob die Ziele und politischen Entscheidungen im Aktionsplan angemessen formuliert wurden, wurde die wissenschaftliche Literatur über Maßnahmen zur Verhinderung von (gewalttätiger) Radikalisierung geprüft und mit dem Aktionsplan verglichen.

Für weitere Informationen: Hardyns, W., Pauwels, L. und Thys, J. (2020) [Een transversale programmascan van het Vlaamse actieplan ter preventie van gewelddadige radicalisering en polarisering.](#) In Cops, D., Pauwels, L. und Van Alstein, M. (eds) Gewelddadige radicalisering & polarisering: Beleid & preventie in Vlaanderen: Evaluatie en uitdagingen.

Gemischte Methoden (Kombination mehrerer Datentypen)

Die oben beschriebenen Datenerhebungsmethoden können einzeln eingesetzt, aber auch kombiniert werden. Dies wird als "Mixed-Methods"-Ansatz bezeichnet. Der Einsatz gemischter Methoden wird bei der Evaluierung von PVE/CVE/Derad-Initiativen immer beliebter.

Ein Ansatz mit gemischten Methoden hat bei sorgfältiger Planung mehrere Vorteile. Insbesondere kann er die Beschränkungen einzelner Datenerhebungsmethoden ausgleichen und wesentlich reichhaltigere Daten über die Initiative liefern. Die Verwendung gemischter Methoden ist nicht automatisch besser als die Verwendung nur einer Datenerhebungsmethode. Es ist wichtig, sorgfältig darüber nachzudenken, welchen Mehrwert die einzelnen Datenerhebungsmethoden haben werden.





13 BEISPIEL 1 FÜR GEMISCHTE METHODEN

Johns et al. Verwendeten eine gemischte Methodik, um die Auswirkungen eines sportbasierten Mentoring-Programms in Melbourne zu bewerten, das sich mit Fragen der Identität, Zugehörigkeit und kulturellen Isolation junger muslimischer Männer befasste. Bei ihrer Evaluierung führten die Forscher:innen in der zweiten Hälfte des Programms eine teilnehmende Beobachtung durch. Nach Abschluss des Programms wurden mit Hilfe qualitativer Methoden (halbstrukturierte Interviews und Fokusgruppen) und quantitativer Methoden (Ausstiegsbefragungen) Daten von Teilnehmern und Interessenvertreter:innen gesammelt, die später zum Vergleich mit den qualitativen Daten herangezogen wurden.

Für weitere Informationen: Johns, A., Grossman, M., & McDonald, K. (2014). "More Than a Game": The Impact of Sport-Based Youth Mentoring Schemes on Developing Resilience towards Violent Extremism. Soziale Eingliederung 2(2), 57–70.



14 BEISPIEL 2 FÜR GEMISCHTE METHODEN

Akademische Forscher:innen führten eine Evaluierung der Prevent-Strategie durch, die darauf abzielt, die Radikalisierung junger Menschen im Vereinigten Königreich zu verhindern. Die Forscher verwendeten einen Ansatz mit gemischten Methoden, der die folgenden Methoden kombiniert:

- eine systematische Überprüfung der Forschungsliteratur,
- Interviews mit Praktikern an allen 48 Projektstandorten in der Anfangsphase der Projektdurchführung,
- die Erfassung von 194 Tagebuchblättern, um die Projektinterventionen und -aktivitäten auf nationaler Ebene zu erfassen,
- eine Analyse der vierteljährlichen Erträge der Projekte und
- vertiefte Fallstudien an 12 Projektstandorten durchgeführt.

Für weitere Informationen: Hirschfield, A., Christmann, K., Wilcox, A., Rogerson, M., & Sharratt, K. (2012). Prozessevaluation der Prävention von gewalttätigem Extremismus: Programmes for Young People. Youth Justice Board.

7.3 METHODEN DER DATENANALYSE

Nachdem die Daten gesammelt wurden, müssen sie analysiert werden, um Schlussfolgerungen zu ziehen. Zu den Methoden der Datenanalyse gehören Techniken zum Sortieren, Interpretieren und Verarbeiten der erhobenen Daten. Auch hier gibt es eine breite Palette von Möglichkeiten. Welche Optionen am besten geeignet sind, hängt davon ab, welche Art von Daten erhoben wurde, welche Fragen beantwortet werden müssen und über welche Fachkenntnisse der/die Evaluator:in verfügt. Im Folgenden werden einige der am häufigsten verwendeten Datenanalysemethoden vorgestellt.

Qualitative Inhaltsanalyse

Diese Methode wird häufig für die Analyse von Dokumenten sowie von Protokollen von Interviews und Fokusgruppen verwendet. Sie kann auch für die Analyse von Notizen aus teilnehmender/nicht teilnehmender Beobachtung verwendet werden.



Bei dieser Methode geht es im Wesentlichen darum, den Text zu strukturieren und zu organisieren, indem die für die Evaluierung relevanten Teile identifiziert und mit bestimmten Themen oder Schlüsselwörtern kodiert werden. Auf diese Weise wird es möglich, alle Teile der Daten herauszufiltern, die sich mit einem bestimmten Thema befassen. Wenn die Daten durch die Kodierung strukturiert und organisiert sind, wird es einfacher, die Daten zu analysieren und zu bestimmen, welche Art von Schlussfolgerungen daraus gezogen werden können.

Die zur Organisation der Daten verwendeten Codes werden von dem/der Evaluator:in entwickelt. Es gibt zwei allgemeine Strategien, wie dies geschehen kann. Deduktives Kodieren ist eine Top-Down-Strategie, bei der die Codes auf der Grundlage der Evaluationsziele und der Schlüsselfragen entwickelt werden. Bei der induktiven Kodierung handelt es sich um eine Bottom-up-Strategie, bei der die Codes auf der Grundlage einer genauen Lektüre der Daten und der Identifizierung der darin enthaltenen Themen und Fragestellungen entwickelt werden. In der Regel werden beide Strategien zusammen angewendet.

Die Kodierung ist vor allem dann hilfreich, wenn viele Textdaten vorhanden sind. Es gibt verschiedene Arten von Software, die für die Kodierung verwendet werden können (z. B. Nvivo und Atlas.ti).

Quantitative Analyse

Quantitative (numerische) Daten, die durch eine Umfrage oder die Überprüfung von Projektdaten gesammelt und erstellt werden können, lassen sich mit verschiedenen statistischen Methoden analysieren. Zu diesen Methoden gehören grundlegende deskriptive Statistiken, die darauf abzielen, einen Datensatz zu beschreiben, z. B. durch die Berechnung der Spannweite und des Durchschnittswerts (Mittelwert, Modus oder Medianwert). Deskriptive Statistiken können mit Softwareprogrammen wie Excel oder SPSS durchgeführt werden. Sie kann als erster Schritt zu einer komplexeren Analyse des Datensatzes dienen, die die Erstellung eines linearen Regressionsmodells beinhalten kann. Ein lineares Regressionsmodell analysiert die statistische Signifikanz der Korrelation und Beziehung zwischen zwei Variablen, die in einem Diagramm dargestellt werden können. Dieses Modell kann ein leistungsfähiges Instrument zur Evaluierung der Auswirkungen einer Initiative sein, vorausgesetzt, es stützt sich auf einen sorgfältig erhobenen Datensatz.

Mehr erfahren

- Hofman, J. und Sutherland, A. (2017). Verschiedene Evaluierungsdesigns und -methoden: Evaluierung von Interventionen zur Prävention oder Bekämpfung von gewalttätigem Extremismus. RAND Europe.
- Wirkung Europa. Datenbank für [Datenerhebungsmethoden und Evaluierungskonzepte](#).
- Strauss, A. (1987). Qualitative Analyse für Sozialwissenschaftler. Cambridge: Cambridge University Press.
- Saldana, J. (2013). The Coding Manual for Qualitative Researchers. London: Sage Publications.
- Peersman, G. (2014) Overview: Data Collection and Analysis Methods in Impact Evaluation, Methodological Briefs: Impact Evaluation 10. Florence: UNICEF Office of Research.
- Trochum, W. (2006). Forschungsmethoden Wissensbasis. [Deskriptive Statistik](#).
- Woodley, A., (2004). [Erfassung und Analyse quantitativer Daten](#). Die PREST-Schulungsressourcen. Commonwealth of Learning.

8. MENSCHENRECHTSBASIERTER ANSATZ DER EVALUIERUNG

Ein wichtiger Teil jeder Evaluierung ist es, sicherzustellen, dass sie ethischen und rechtlichen Standards entspricht. Dies gilt umso mehr für PVE/CVE/Derad-Initiativen, die sich mit sensiblen Themen befassen und sich häufig an gefährdete und stigmatisierte Personen und Gemeinschaften richten. Eine Evaluierung kann auch eine Möglichkeit sein, um zu beurteilen, inwieweit die Initiative selbst ethischen und rechtlichen Standards entspricht.

Ein guter Rahmen für ethische Überlegungen bei PVE/CVE/Derad-Initiativen und -Evaluierungen ist der auf den Menschenrechten basierende Ansatz. In diesem Kapitel werden kurz die Grundlagen dieses Rahmens und seine Bedeutung für die Praxis vorgestellt. Detaillierte Anleitungen zur angemessenen Einbeziehung ethischer, geschlechtsspezifischer, rechtlicher und gesellschaftlicher Aspekte in die Evaluierung finden Sie im INDEED-Guide 2.

Menschenrechtsbasierter Ansatz

Der menschenrechtsbasierte Ansatz (Human Rights-Based Approach, HRBA) ist ein konzeptioneller Rahmen, der auf jede Politik oder Praxis angewendet werden kann, um sicherzustellen, dass sie normativ auf internen Menschenrechtsstandards basiert und operativ auf den Schutz der Menschenrechte ausgerichtet ist.

Der HRBA stützt sich auf fünf zentrale Menschenrechtsgrundsätze:

Teilnahme	Jeder Mensch hat das Recht auf aktive Beteiligung an Entscheidungsprozessen, die sich auf die Wahrnehmung seiner/ihrer Rechte auswirken.
Rechenschaftspflicht	Praktiker:innen und Manager:innen von PVE/CVE/Derad-Initiativen und Evaluierungen werden zur Rechenschaft gezogen, wenn sie ihren Verpflichtungen gegenüber den Zielgruppen nicht nachkommen. Bei Menschenrechtsverletzungen sollten wirksame Rechtsmittel zur Verfügung stehen.
Anti-Diskriminierung und Gleichstellung	Alle Menschen haben Anspruch auf ihre Rechte ohne jede Art von Diskriminierung. Alle Arten von Diskriminierung sollten verboten, verhindert und beseitigt werden.
Ermächtigung	Jeder hat das Recht, seine Rechte einzufordern und auszuüben. Einzelpersonen und Gemeinschaften müssen ihre Rechte kennen und an der Entwicklung von Maßnahmen, die ihr Leben betreffen, mitwirken.
Rechtmäßigkeit	Die Initiativen und ihre Evaluierungen sollten mit den im nationalen und internationalen Recht verankerten Rechten in Einklang stehen.

Nachzulesen unter European Network of National Human Rights Institutions, [Human Rights-Based Approach](#).

Was bedeutet die Umsetzung dieser Werte und Grundsätze in PVE/CVE/Derad-Initiativen und Evaluierungen? Im Rahmen des INDEED-Projekts haben wir die Schlüsselfragen in vier Dimensionen gegliedert: geschlechtsspezifische Aspekte, ethische Aspekte, rechtliche Aspekte und gesellschaftliche Aspekte (GELSA).

Geschlechtsspezifische Aspekte

Es wird empfohlen, bei der Evaluierung von PVE/CVE/Derad einen geschlechtsspezifischen Ansatz zu wählen. Dies gilt insbesondere dann, wenn die zu bewertende Initiative in ihrer Konzeption nicht geschlechtersensibel ist. Radikalisierungsprozesse sind nicht geschlechtsneutral, und dies gilt auch für PVE/CVE/Derad-Initiativen, selbst wenn sie sich nicht explizit mit dem Thema Geschlecht befassen. Dies liegt daran, dass das Geschlecht unvermeidlich einen Einfluss auf die Erfahrungen und Bedingungen des Einzelnen hat. Es ist auch bekannt, dass es geschlechtsspezifische Unterschiede in Radikalisierungsprozessen geben kann.

Ein geschlechtersensibler Evaluierungsansatz berücksichtigt die potenziellen geschlechtsspezifischen Unterschiede und ist so konzipiert, dass er diese auch erfassen kann. Dies kann zum Beispiel bedeuten, dass sichergestellt wird, dass ein ausgewogenes Verhältnis zwischen den Geschlechtern unter den Befragten und in den Fokusgruppen besteht und dass die gesammelten Daten auch aus der Geschlechterperspektive analysiert werden, wobei die Auswirkungen auf Menschen aller Geschlechter berücksichtigt und bewertet werden.

Ethische Aspekte

PVE/CVE/Derad-Initiativen und ihre Evaluierungen haben oft mit Personen und Gruppen zu tun, die stigmatisiert sind und kontroverse politische oder religiöse Ansichten vertreten. Es ist wichtig, sicherzustellen, dass Initiativen in diesem Bereich in ihren Praktiken oder Annahmen nicht diskriminierend sind. Es gibt negative Beispiele für PVE/CVE/Derad-Initiativen in der EU, bei denen sich Einzelpersonen und Gemeinschaften aufgrund ihrer religiösen, ethnischen oder sozialen Identität angegriffen fühlten und die Initiative eine Erweiterung gesellschaftlicher Vorurteile darstellte.

Eine Evaluierung bietet eine gute Gelegenheit, die zugrundeliegenden und offenkundigen Annahmen der PVE/CVE/Derad-Initiative in Bezug auf ihre Zielgruppen sowie die Art und Weise, wie sich diese Annahmen in den Zielen, Leitlinien und Praktiken der Initiative manifestieren, sorgfältig zu bewerten.

Eine Evaluierung befasst sich auch mit der Frage, wie die Initiative die verschiedenen Freiheiten berücksichtigt hat, die mit dem Recht auf Nichtdiskriminierung verbunden sind, wie etwa die Gedanken-, Meinungs-, Versammlungs- und Vereinigungsfreiheit. Im Rahmen einer Evaluierung kann beurteilt werden, wie diese Rechte und Freiheiten mit den Erfordernissen der PVE/CVE/Derad-Arbeit zur Bekämpfung von Radikalisierung und Extremismus in Einklang gebracht wurden und wie gut es der Initiative gelungen ist, die Verstärkung von Ausgrenzungsmechanismen zu vermeiden, die Menschen dem Risiko der Radikalisierung aussetzen.

Gute Leitlinien für die ethische Durchführung einer Evaluierung beruhen auch auf den allgemeinen Grundsätzen einer verantwortungsvollen Forschung. Die Ethikrichtlinien für die Forschung enthalten gute Anleitungen für den Umgang mit Befragten und Fokusgruppen und für das Verfassen eines Evaluierungsberichts, der deren Rechte und Integrität berücksichtigt.

Rechtliche Aspekte

Neben den bereits erörterten Grundrechtsbestimmungen müssen auch die nationalen Rechtsvorschriften berücksichtigt werden. Auch sektorspezifische berufsständische Regelungen können sich beispielsweise auf die Möglichkeiten der Datenerhebung und -weitergabe auswirken. Daher muss sichergestellt werden, dass sowohl die Initiative als auch ihre Evaluierung auf einer angemessenen Rechtsgrundlage beruhen.

Ein zentrales Thema ist, dass fast alle Initiativen und Evaluierungen auf Bedenken hinsichtlich personenbezogener Daten stoßen. Die [Allgemeine Datenschutzverordnung](#) (DSGVO) legt fest, wie personenbezogene Daten verarbeitet (d.h. erhoben, gespeichert, analysiert oder weitergegeben) werden dürfen. Die Verarbeitung personenbezogener Daten erfordert immer eine spezifische Rechtsgrundlage, wenn sie nicht durch die ausdrückliche und in Kenntnis der Sachlage erteilte Einwilligung der betreffenden Person erlaubt ist.

Neben der Einhaltung der DSGVO-Vorschriften und anderer Gesetze können PVE/CVE/Derad-Evaluierungen die Initiative auch dahingehend beurteilen, ob ihre Praktiken mit diesen Vorschriften in Einklang stehen. Es ist wichtig, dass diese Vorschriften genau befolgt werden, da Datenverstöße, -missbrauch und -verletzungen Einzelpersonen gefährden und das Vertrauen in PVE/CVE/Derad-Praktiker:innen und ihre Arbeit zerstören können. Solche Verstöße können auch Ängste vor Überwachung, Profiling und vorurteilsbehafteten Praktiken verstärken.

Gesellschaftliche Aspekte

Eine Evaluierung kann auch die Gelegenheit bieten, über die eigentliche Initiative hinaus zu blicken und zu bewerten, wie gut sie auf den breiteren gesellschaftlichen Kontext eingeht, in dem sie agiert, und welche Art von Auswirkungen sie in Bezug auf diesen Kontext hat. Diese Art von Sensibilität für den gesellschaftlichen Kontext ist oft ein Schlüssel für den Erfolg von PVE/CVE/Derad-Maßnahmen und die Vermeidung unbeabsichtigter schädlicher Auswirkungen.

Die Evaluierung kann sich auf die Frage konzentrieren, wie (und ob) die Initiative das Risiko unbeabsichtigter Folgen ihrer Arbeit für die Gemeinschaften und die Gesellschaft angegangen ist, wie z. B. Stigmatisierung, Eingliederung/Ausgrenzung, vorherrschende Vorurteile und Stereotypen sowie das Gefühl der Sicherheit innerhalb bestimmter Gemeinschaften und der Gesellschaft im Allgemeinen.

Mehr erfahren

- Europäisches Netzwerk der nationalen Menschenrechtsinstitutionen, [Menschenrechtsbasierter Ansatz](#).
- [Charta der Grundrechte der Europäischen Union](#).
- RAN. [Geschlechtsspezifische Ansätze in der PVE: Präventive Arbeit für Mädchen in und um Schulen](#). Abschlusspapier der RAN-Arbeitsgruppensitzung.
- Lloyd, M. [Ethical Guidelines for Working on P/CVE in Mental Health Care](#). RAN-Veröffentlichung.
- Evaluierungsgruppe der Vereinten Nationen. [Ethische Richtlinien für Evaluierung](#).

9. ACHT TIPPS FÜR ERFOLGREICHE EVIDENZBASIERTE EVALUATIONEN

Frühere Studien und Evaluierungen haben gezeigt, dass sich bei PVE/CVE/Derad-Evaluierungen häufig bestimmte Herausforderungen ergeben haben. Im Folgenden haben wir eine Liste von Tipps zusammengestellt, wie diese vermieden oder gemildert werden können:

1

1. Integration der Evaluierung in den Initiativplan

Viele Herausforderungen bei der Evaluierung können vermieden werden, wenn die Evaluierung zusammen mit der Initiative selbst geplant wird. Das bedeutet, dass im Plan der Initiative auch festgelegt wird, wie sie bewertet werden soll. Die erste Evaluierung kann bereits in dieser Planungsphase stattfinden. Viele Entscheidungen, die in den ersten Tagen der Laufzeit der Initiative getroffen werden, bestimmen, welche Art von Evaluierungsdesigns, -methoden und -daten in späteren Evaluierungen verwendet werden können.

2

2. Finanzierung und Ressourcen für die Evaluierung sichern

Eine häufige Herausforderung für die Evaluierung besteht darin, dass es dafür keine Mittel gibt. Dies scheint besonders häufig bei langfristigen Programmen und kurzfristigen Maßnahmen der Fall zu sein, die auf eine befristete Projektfinanzierung angewiesen sind. In diesen Fällen kann der Geldgeber verlangen, dass die Initiative evaluiert wird, aber es gibt keine zusätzlichen Mittel für diesen Zweck.

Eine Evaluierung, selbst in ihrer einfachsten Form, erfordert beträchtliche Ressourcen. Es sollte klar geplant werden, woher diese Ressourcen kommen sollen. Es wird empfohlen, für die Finanzierung von PVE/CVE/Derad-Initiativen zusätzlich zu den Mitteln für die Durchführung der Initiative selbst ein eigenes Budget für die Evaluierung bereitzustellen.

3

3. Definieren Sie die Konzeption und die Ziele der Initiative klar und deutlich

Ausgangspunkt für praktisch jede Art von Evaluierung ist ein klares Verständnis der Ziele, des Umsetzungsplans und der zugrunde liegenden Annahmen darüber, wie die geplanten Maßnahmen zu den beabsichtigten Ergebnissen führen sollen (Theory of Change). Im Idealfall werden die Ziele und die Theory of Change in der Planungsphase der Initiative sorgfältig entwickelt. Wenn die Ziele nicht richtig definiert wurden, ist es schwierig, eine qualitativ hochwertige Evaluierung durchzuführen.



4. Entwicklung der Datenüberwachungspraktiken der Initiative im Hinblick auf die Evaluierung

Die Art der Evaluierung, die durchgeführt werden kann, hängt stark davon ab, welche Daten zur Verfügung stehen. Es gibt viele Arten von Daten, die während der Evaluierung erhoben werden können, aber einige Evaluierungskonzepte erfordern Daten, die nur zu einem bestimmten Zeitpunkt erhoben werden können. Um beispielsweise die Auswirkungen einer Initiative auf die Teilnehmer zuverlässig bewerten zu können, sollten einige Informationen über ihre Situation zu Beginn der Maßnahme vorliegen. Die Evaluierung erfordert fast immer eine gute Dokumentation der Aktivitäten der Initiative.

Wenn die Initiative über systematische Überwachungs- und Aufzeichnungspraktiken verfügt, die sich an den Erfordernissen künftiger Evaluierungen orientieren, werden die erforderlichen Daten bei Bedarf zur Verfügung stehen. Unzureichende Überwachungspraktiken machen eine Evaluierung nicht unmöglich, aber sie schränken die verfügbaren Optionen erheblich ein und können sich negativ auf die Zuverlässigkeit der Ergebnisse auswirken.

5. Fokus auf Lernen

Eine Evaluierung ist am nützlichsten, wenn sie zu Lernzwecken geplant und durchgeführt wird. Das bedeutet, dass die Evaluierung so konzipiert ist, dass sie die weitere Entwicklung der Initiative unterstützt, und dass sie einen klaren Plan für die Verwendung der Ergebnisse enthält. Es wird empfohlen, die Evaluierung nicht an Finanzierungsentscheidungen zu knüpfen, sondern sie in der Mitte eines Finanzierungszeitraums durchzuführen. Dies bietet das konstruktivste Umfeld für die Durchführung der Evaluierung ohne Erfolgsdruck und bietet eine echte Gelegenheit, die Ergebnisse für die Verbesserung der Initiative zu nutzen.

6. Einbindung der wichtigsten Interessengruppen in die Planung der Evaluierung

Evaluierungen sind oft ein Top-down-Prozess, bei dem die Zielsetzung und das Design von dem/der Geldgeber:in festgelegt werden. Die Erfahrung hat gezeigt, dass dies keine sehr produktive Strategie ist. Stattdessen ist es in der Regel besser, einen partizipatorischen Ansatz zu verfolgen und die wichtigsten Interessengruppen, insbesondere diejenigen, die die Initiative durchführen, von Anfang an in den Evaluierungsprozess einzubeziehen.

Die Einbeziehung von Interessengruppen ermöglicht es, die Eigenverantwortung und das Vertrauen zu schaffen, die notwendig sind, um zuverlässige Ergebnisse zu erzielen. Wenn die Beteiligten aktiv in den Planungsprozess einbezogen werden, zeigt dies, dass ihre Ansichten, ihr Wissen und ihre Erfahrungen geschätzt werden.

Wenn die Evaluierung auf die Bedürfnisse der an der Umsetzung des Vorhabens Beteiligten eingeht, wird die Evaluierung nicht mehr als eine zusätzliche Aufgabe betrachtet, die von jemand anderem verlangt wird, sondern als etwas, das für die Entwicklung der eigenen Arbeit von Nutzen ist. Dies ist ein weiterer Anreiz, bei der Erstellung der für die Evaluierung benötigten Daten mitzuarbeiten und ehrlich über mögliche Schwächen und Fehler zu sprechen.

7

7. Vergewissern Sie sich, dass der/die Evaluator:in ein gutes Verständnis vom Bereich PVE/CVE/Derad hat

Wenn die Evaluierung von einer:m externen Evaluator:in durchgeführt wird, ist es gut sicherzustellen, dass der/die Evaluator:in mit dem PVE/CVE/Derad-Bereich und seinen spezifischen Merkmalen vertraut ist. Zwar können viele gängige Evaluierungsarten und -methoden auf die Evaluierung von PVE/CVE/Derad-Initiativen angewandt werden, doch ist es ziemlich schwierig, Daten oder Beobachtungen ohne ein gutes allgemeines Verständnis des Bereichs sinnvoll zu nutzen.

8

8. Frühzeitige Überlegungen zu ethischen Fragen und zur Sicherstellung der Datenverfügbarkeit

Eine häufige Herausforderung bei der Evaluierung von PVE/CVE/Derad-Initiativen besteht darin, dass ethische und Sicherheitsbedenken den Zugang zu Daten einschränken können. Regierungsbehörden und diejenigen, die die Initiative durchführen, können zögern, Informationen weiterzugeben, insbesondere wenn sie Details über die politischen Meinungen oder persönlichen Angelegenheiten einer Person enthalten. Sicherheitsaspekte können den Zugang zu Orten einschränken, an denen Interviews oder Beobachtungen durchgeführt werden könnten. Es empfiehlt sich, frühzeitig herauszufinden, welche ethischen, sicherheitstechnischen und rechtlichen Einschränkungen für die gemeinsame Nutzung von Daten bestehen, und zu überlegen, wie innerhalb dieser Grenzen die bestmöglichen Ergebnisse erzielt werden können.

Eine weitere Reihe von Herausforderungen ergibt sich aus dem sensiblen und stigmatisierten Charakter des gewalttätigen Extremismus. Es kann schwierig sein, die Zielgruppen und Einzelpersonen einer Initiative zu erreichen und Vertrauen zu ihnen aufzubauen, da die Situation und die Beziehungen zu Personen, die in den Extremismus verwickelt sind, fragil sind. Möglicherweise sind sie auch nicht bereit, ihre wahren Ansichten mitzuteilen, da diese als gesellschaftlich unerwünscht und stigmatisiert angesehen werden könnten. Daher ist es ratsam, sich genau zu überlegen, wen man als Evaluator:in auswählt, um das für die Datenerhebung erforderliche Vertrauen zu gewinnen – und zwar auf eine Weise, die das Vertrauen der Personen in die Initiative selbst nicht gefährdet.

QUELLEN UND LITERATUR

- Aaltonen, M. & Hinkkanen, V. (2014). Katuväkivaltatyön vaikuttavuus: HelsinkiMission Aggredi-ohjelman arviointi. Verkkokatsauksia 40/2014. Oikeuspoliittinen tutkimuslaitos. https://helda.helsinki.fi/bitstream/handle/10138/152599/40_HelsinkiMission.pdf?sequence=2&isAllowed=y.
- Aldrich, D.P. (2014). First Steps Towards Hearts and Minds? USAID's Countering Violent Extremism Policies in Africa. *Terrorism and Political Violence*, 26(3), 523–46.
- Bellasio, J., Hofman, J., Ward, A., Nederveen, F., Knack, A., Meranto, A. S., & Hoorens, S. (2018). Counterterrorism evaluation: Taking stock and look ahead. California and UK: RAND Europe research report. https://www.rand.org/pubs/research_reports/RR2628.html.
- BetterEvaluation. Realist Evaluation. <https://www.betterevaluation.org/methods-approaches/approaches/realist-evaluation>.
- BetterEvaluation. Utilisation-focused evaluation. <https://www.betterevaluation.org/methods-approaches/approaches/utilisation-focused-evaluation>.
- BetterEvaluation. What is evaluation? <https://www.betterevaluation.org/getting-started/what-evaluation>.
- Bjørge, T. (2013). Strategies for preventing terrorism. Hampshire, UK: Palgrave Pivot.
- Bourekba, M. (2021). Detect to prevent: countering violent extremism strategies in Spain. *Revista CIDOB d'Afers Internacionals*, 128, 81–105. doi.org/10.24241/rci.2021.128.2.81/en
- Breuer, E., De Silva, M. J., Fekadu, A., Luitel, N. P., Murhar, V., Nakku, J., & Lund, C. (2014). Using workshops to develop theories of change in five low and middle income countries: Lessons from the programme for improving mental health care (PRIME). *International journal of mental health systems*, 8(1).
- Brophy, S., Snooks, H., & Griffiths, L. (2008). Small-scale evaluation in health: A practical guide. London: Sage.
- Busher, J., Malkki, L., & Marsden, S. (eds.) (2023). *Routledge Handbook on Radicalisation and Countering Radicalisation*. Routledge.
- Byrne-Diakun, R.M. (Ed.). (2016). *Countering Violent Extremism: Applying the Public Health Model*. Georgetown University Center for Security Studies. National Security Critical Issues Task Force. <https://georgetownsecuritystudiesreview.org/wp-content/uploads/2016/10/NSCITF-Report-on-Countering-Violent-Extremism.pdf>.
- Carson, J.V. (2017). Assessing the Effectiveness of High-Profile Targeted Killings in the “War on Terror”: A Quasi-Experiment. *Criminology and Public Policy*, 16(1), 191–220.
- Charter of Fundamental Rights of the European Union. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT>.
- Cherney, A., & Belton, E. (2020). Assessing intervention outcomes targeting radicalised offenders: Testing the pro integration model of extremist disengagement as an evaluation tool. *Dynamics of Asymmetric Conflict*, 13(3), 193–211.
- Cherney, A., Belton, E., & Koehler, D. (2020). Program integrity guidelines to inform the design, implementation, and evaluation of P/CVE initiatives: Guideline review document & Program integrity tool. Countering Violent Extremism Unit, Department of Justice and Community Safety.
- Chowdhury Fink, N., Romaniuk, P. & Barakat, R. (2013). Evaluating Countering Violent Extremism Programming: Practice and progress. Center on Global Counterterrorism Cooperation.
- Clark, H., & Anderson, A. A. (2004). Theories of change and logic models: Telling them apart. American Evaluation Association Conference. https://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/TOCs_and_Logic_Models_forAEA.pdf.
- Cox, K., & Van Gorp, A. (2018). Surveys. In J. Hofman and A. Sutherland (eds). *Evaluating interventions that prevent or counter violent extremism*. RAND Europe. https://www.rand.org/pubs/research_reports/RR2094.html.
- CREST. Introductory Guide: Countering Violent Extremism. CREST: Centre for Research and Evidence on Security Threats. https://eprints.lancs.ac.uk/id/eprint/88097/1/17_008_01.pdf.
- Davies, P. (2012). The State of Evidence-Based Policy Evaluation and its Role in Policy Formation. *National Institute Economic Review*, 219(1), R41–R52.
- Dhillon, L. & Vaca, S. (2018). Refining theories of change. *Evaluation*, 14(30), 64–87.
- Dunn, K. M., Atie, R., Kennedy, M., Ali, J. A., O'Reilly, J., & Rogerson, L. (2016). Can you use community policing for counter terrorism? Evidence from NSW, Australia. *Police Practice and Research*, 17(3), 196–211.
- EUCPN (2013). Evaluation of crime prevention initiatives: the principles of evaluation. In EUCPN Secretariat (eds.), *EUCPN Thematic Paper Series*, no. 5, European Crime Prevention Network: Brussels.
- European Commission (2014). Preventing Radicalisation to Terrorism and Violent Extremism: Strengthening the EU's Response. Brussels, Belgium: European Commission. [https://www.europarl.europa.eu/meetdocs/2009_2014/documents/com/com_com\(2013\)0941/com_com\(2013\)0941_en.pdf](https://www.europarl.europa.eu/meetdocs/2009_2014/documents/com/com_com(2013)0941/com_com(2013)0941_en.pdf).
- European Commission (2021). Better Regulation Guidelines. European Commission. https://commission.europa.eu/system/files/2021-11/swd2021_305_en.pdf.
- European Network of National Human Rights Institutions, Human Rights-Based Approach, <https://ennhri.org/>



about-nhris/human-rights-based-approach/.

- Evaluation toolbox. Types of evaluation. https://evaluationtoolbox.net.au/index.php?option=com_content&view=article&id=15&Itemid=19.
- Feddes, A. R., Mann, L., & Doosje, B. (2015). Increasing Self-esteem and Empathy to Prevent Violent Radicalization: A Longitudinal Quantitative Evaluation of a Resilience Training Focused on Adolescents with a Dual Identity. *Journal of Applied Social Psychology*, 45(7), 400-411.
- Fish, D., & Coles, C. (1998). Developing professional judgement in health care: learning through the critical appreciation of practice. Oxford: Butterworth Heinemann.
- Fisher, T., Range, D., & Cuddihy, J. (2020). Evaluation of 'Violent Extremism Strengthening Resilience (STRIVE II) in Kenya: Final report. <https://ct-morse.eu/resource/evaluation-of-strive-ii-in-kenya/>.
- Gielen, A.J. (2017). Evaluating Countering Violent Extremism. In L. Colaert (Ed.). *Deradicalisation. Scientific insights for policy*. Brussels: Flemish Peace Institute.
- Gielen, A.J. (2019). Countering Violent Extremism: A Realist Review for Assessing What Works, for Whom, in What Circumstances, and How? *Terrorism and Political Violence*, 31(6), 1149-1167.
- Hardyns, W., Pauwels, L. and Thys, J. (2020). Een transversale programmascan van het Vlaamse actieplan ter preventie van gewelddadige radicalisering en polarisering. In D. Cops, L. Pauwels, and M. Van Alstein (eds). *Gewelddadige radicalisering & polarisering: Beleid & preventie in Vlaanderen: Evaluatie en uitdagingen*. <https://vlaamsvredesinstituut.eu/rapport/gewelddadige-radicalisering-en-polarisering-beleid-preventie-in-vlaanderen-evaluatie-en-uitdagingen/>.
- Hardyns, W., Klima, N., & Pauwels, L. (Eds.). (2022). *Evaluation and mentoring of the multi-agency approach to violent radicalisation in Belgium, the Netherlands and Germany*. Antwerpen: Maklu.
- Harris-Hogan, S., Barrelle, K., & Zammit, A. (2016). What Is Countering Violent Extremism? Exploring CVE Policy and Practice in Australia. *Behavioral Sciences of Terrorism and Political Aggression*, 8(1), 6-24.
- Hirschfield, A., Christmann, K., Wilcox, A., Rogerson, M., & Sharratt, K. (2012). *Process Evaluation of Preventing Violent Extremism Programmes for Young People*. London: Youth Justice Board.
- Hill, J. R., & Thies, J. (2010). Program theory and logic model to address the co-occurrence of domestic violence and child maltreatment. *Evaluation and program planning*, 33(4), 356-364.
- Hofman, J., & Sutherland, A. (eds.) (2018). *Evaluating interventions that prevent or counter violent extremism*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR2094.html.
- Holmer, G., Bauman, P. & Aryaeinejad, K. (2018). *Measuring Up: Evaluating the impact of PVE/CVE/DERAD Programs*. USA: USIP. <https://www.usip.org/sites/default/files/2018-09/preventing-countering-violent-extremismmeasuringup.pdf>.
- Iacopini, G., Stock, L. & Junge, K. (2011). *Evaluation of Tower Hamlets Prevent Projects Final report prepared for the London Borough of Tower Hamlets*. The Tavistock Institute. http://www.tavinstitute.org/wp-content/uploads/2012/12/Tavistock_Projects_Tower-Hamlets-PVE_evaluation_final_report.pdf.
- Impact Europe. Database of data collection methods and evaluation designs. <http://www.impact.itti.com.pl/index#/guide/repository>.
- James Bell Associates. (2018). *Formative evaluation toolkit: A step-by-step guide and resources for evaluating program implementation and early outcomes*. Washington, DC: Children's Bureau, Administration for Children and Families, U.S. Department of Health and Human Services. https://www.acf.hhs.gov/sites/default/files/documents/cb/formative_evaluation_toolkit.pdf.
- Janta, B., & Van Gorp, A. (2018). Interviews. In J. Hofman, & A. Sutherland (eds). *Evaluating interventions that prevent or counter violent extremism*. RAND Europe. https://www.rand.org/pubs/research_reports/RR2094.html.
- Johns, A., Grossman, M., & McDonald, K. (2014). "More Than a Game": The Impact of Sport-Based Youth Mentoring Schemes on Developing Resilience toward Violent Extremism. *Social Inclusion* 2(2), 57-70.
- Katz J., Wandersman A., Goodman R.M., Griffin S., Wilson D.K., & Schillaci M. (2013). Updating the FORECAST formative evaluation approach and some implications for ameliorating theory failure, implementation failure, and evaluation failure. *Eval Program Plann*, 39, 42-50.
- Koehler, D. (2017). *Understanding Deradicalization: Methods, Tools and Programs for Countering Violent Extremism*. Routledge.
- Korn, J. (2016). European CVE Strategies from a Practitioner's Perspective. *The ANNALS of the American Academy of Political and Social Science*, 668(1), 180-197.
- Kurtz, J., Wolfe, R., & Tesfaye, B. (2016). Does Youth Employment Build Stability? Evidence From an Impact Evaluation of Vocational Training in Afghanistan. In S. Zeiger (Ed.) *Expanding Research on Countering Violent Extremism*. Abu Dhabi: Hedayah.
- Lipsey, M., Petrie, C., Weisburd, D., & Gottfredson, D. (2006). Improving evaluation of anti-crime programs:

Summary of a National Research Council report. *Journal of Experimental Criminology*, 2(3), 271–307.

- Lloyd, M. Ethical Guidelines for Working on P/CVE in Mental Health Care. RAN publication. https://home-affairs.ec.europa.eu/system/files/2021-04/ran_ethical_guidelines_for_working_p-cve_mhc_2021_en.pdf.
- Madriaza, P., Ponsot, A.S., Marion, D. & Monnier, C. (2017). The prevention of radicalisation leading to violence: An international study of front-line workers and intervention issues. International Center for the Prevention of Crime.
- Madriaza, P., Morin, D., Ousman, S., Autixier, C., Hassan, G., & Venkatesh, V. (2022). Improving evaluations of programs for prevention of radicalization and violent extremism: An exploratory international study. UNESCO chair in Prevention of Radicalisation and Violent Extremism (UNESCO-PREV Chair) in collaboration with the International Centre for the Prevention of Crime, University of Sherbrooke.
- McDavid, J. C., Huse, I., & Hawthorn, L. R. (2019). Program evaluation and performance measurement: an introduction to practice – third edition. SAGE Publications, Inc.
- MEASURE Evaluation. Economic Evaluation. <https://www.measureevaluation.org/our-work/economic-evaluation.html>.
- Minor, D. (2016). MAXIME Berlin – Interkulturelles und interreligiöses Präventionsprojekt zur Toleranz und Demokratieentwicklung Evaluation 2014 – 2016. Berlin: Violence Prevention Network.
- Molenkamp, M., Wouterse, L., & Gielen, A. J. (2018). Ex post paper guidelines evaluations of PCVE programmes and interventions. RAN center for excellence. https://home-affairs.ec.europa.eu/system/files/2020-09/ms_workshops_guidelines_evaluation_of_pcve_programmes_and_interventions_july_2018_en.pdf.
- Morenike Latifat Moroof, M. (2016). An Analysis of the 2016 Gothenburg CVE Framework. Doctoral thesis. Duke Office of Research Support.
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W. et al. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;250:h1258.
- Nehlsen, I., Biene, J., Coester, M., Greuel, F., Milbradt, B., & Armbrorst, A. (2020). Evident and effective? The challenges, potentials and limitations of evaluation research on preventing violent extremism. *International Journal of Conflict and Violence*, 14(2), 1–20.
- Neumann, P. R. (2017). Countering Violent Extremism and Radicalisation That Lead to Terrorism: Ideas, Recommendations, and Good Practices from the OSCE Region. Vienna: OSCE.
- Noordegraaf, M., Douglas, S., Bos, A., & Klem, W. (2016). Gericht, gedragen en geborgd interventievermogen? Evaluatie van de nationale contraterrorisme-strategie 2011–2015. Den Haag: WODC.
- Parker, D. & Lindekilde, L. (2020). Preventing Extremism with Extremists: A Double-Edged Sword? An Analysis of the Impact of Using Former Extremists in Danish Schools. *Education Science*, 10(4), 1–19.
- Patton, M. Q. (1997). Utilization-focused evaluation: The new century text (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2008). Utilization-focused evaluation (4th. Ed). Thousand Oaks, CA: Sage.
- Peersman, G. (2014). Overview: Data Collection and Analysis Methods in Impact Evaluation, Methodological Briefs: Impact Evaluation 10, Florence: UNICEF Office of Research.
- Polish Platform for Homeland Security. Understand = Respect. <https://ppbw.pl/en/preventiveprogram/>.
- RAN (2022). RAN Gender-Specific Approaches in PVE: Preventive Work for Girls in and around Schools. Conclusion paper, RAN working group meeting 31 May–1 June 2022. https://home-affairs.ec.europa.eu/system/files/2022-08/ran_y-e_gender-specific_approaches_in_pve_lisbon_3105-01062022_en.pdf.
- Ris, L. & Ernstofer, A. (2017). Borrowing a Wheel: Applying Existing Design, Monitoring and Evaluation Strategies to Emerging Programming Approaches to Prevent and Counter Violent Extremism. Peacebuilding Evaluation Consortium.
- Robson, C. (2000). Small-scale evaluation: Principles and practice. London: Sage.
- Rogers, P. (2014). Theory of Change. Methodological Briefs: Impact Evaluation 2. UNICEF Office of Research, Florence. https://www.betterevaluation.org/sites/default/files/Theory_of_Change_ENG.pdf.
- Rychetnik, L., Frommer, M., Hawe, P., et al. (2002). Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology & Community Health*, 56, 119–127.
- Rychetnik, L., Hawe, P., Waters, E., Barratt, A., & Frommer, M. (2004). A glossary for evidence based public health. *J Epidemiol Community Health*, 58(7), 538–45.
- Rycroft-Malone, J., Seers, K., Titchen, A., Harvey, G., Kitson, A., & McCormack, B. (2004). What counts as evidence in evidence-based practice?. *Journal of Advanced Nursing*, 47, 81–90.
- Saldana, J. (2013). The Coding Manual for Qualitative Researchers. London. Sage Publications.
- Schmid, A. (2013). Radicalisation, de-radicalisation, counter radicalisation: A conceptual discussion and literature review. The Hague, the Netherlands: International Centre for Counter Terrorism. [ICCT-Schmid-](https://www.icct.nl/publications/ICCT-Schmid-2013)

[Radicalisation-De-Radicalisation-Counter-Radicalisation-March-2013_2.pdf](#)

- Schuurman, B. & Bakker, E. (2016). Reintegrating Jihadist Extremists: Evaluating a Dutch Initiative, 2013-2014. *Behavioural Sciences of Terrorism and Political Aggression*, 8(1), 66-85.
- Stake, R. E. (2004). *Standards-based & responsive evaluation*. SAGE Publications, Inc.
- Stern, E. (2015). *Impact Evaluation: A guide for commissioners and managers*. Bond for International Development.
- Strauss, A. (1987). *Qualitative Analysis for Social Scientists*. Cambridge: Cambridge University Press.
- Thompson, S. & Leroux, E. (2023). Lessons learned from dual site formative evaluations of Countering violent extremism (CVE) programming coiled by Canadian police. *Journal of Policing, Intelligence and Counter Terrorism*, 18(1).
- Trochum, W. (2006). Research Methods Knowledge Base. Descriptive statistics. <https://www.socialresearchmethods.net/kb/statdesc.php>.
- UNISDIR (2021). Monitoring and Evaluation Framework. UNISDIR. https://www.preventionweb.net/files/49324_unisdrmeframeworkver1.0.pdf.
- United Nations Evaluation Group. Ethical Guidelines for Evaluation. <http://www.unevaluation.org/document/download/3625>.
- Veldhuis, T.M., Gordijn, E.H., Lindenberg S. M., & Veenstra, R. (2011). *Terrorists in Prison: Evaluation of the Dutch Terrorism Wing*. Groningen: University of Groningen.
- Vergani, M., Barton, G., Iqbal, M., & Kruber, S. (2016). How to Design Impact Evaluations of CVE programs: a practical guide for Southeast Asian civil society organizations. Sean-Cso.
- Vermeulen, F. (2014). Suspect Communities-Targeting Violent Extremism at the Local Level: Policies of Engagement in Amsterdam, Berlin, and London. *Terrorism and Political Violence*, 26, 286-306.
- Vidino, L. & Seamus, H. (2015). *Countering Violent Extremism in America*. Center for Cyber & Homeland Security-The George Washington Institute.
- Webster, S., Kerr, J. & Tompkins, C. (2017). *A Process Evaluation of the Structured Risk Guidance for Extremist Offenders*. London: HM Prison & Probation Service.
- Weine, S., Eisenman, D., Glik, D., Kinsler, J., & Polutnik, C. (2016). Leveraging a targeted violence prevention program to prevent violent extremism: A formative evaluation in Los Angeles. University of Illinois Chicago and University of California in Los Angeles.
- Weine, S., Horgan, J., Robertson, C., Loue, S., Mohamed, A., & Noor, S. (2009). Community and family approaches to combating the radicalization and recruitment of Somali-American youth and young adults: A psychosocial perspective. *Dynamics of Asymmetric Conflict*, 2, 181-200.
- WHO. Economic Evaluations. https://www.unodc.org/docs/treatment/economic_evaluation.pdf.
- Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2004). *Handbook of practical program evaluation* (2nd ed.). San Francisco: Jossey-Bass.
- Williams, M.J. (2021). *Preventing and Countering Violent Extremism: Designing and Evaluating Evidence-Based Programs*. Routledge.
- Wong-Reiger, D. (1993). Introduction to the Program Logic Model. *A Hands-On Guide to Planning and Evaluation*. Canadian Hemophilia Society, 23-32.
- Woodley, A. (2004). Getting and analysing of quantitative data. The PREST training resources. Commonwealth of Learning. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.8332&rep=rep1&type=pdf>.
- Youker, B. W., & Ingraham, A. (2014). Goal-Free Evaluation: An Orientation for Foundations' Evaluations. *The Foundation Review*, 5(4).
- Zimmerman, M., & Holden, D. (Eds.) (2009). *Program evaluation planning: Overview and analysis*. SAGE Publications, Inc.





INDEED

Evidence-Based Model for Evaluation of
Radicalisation Prevention and Mitigation