

## Evaluationskapazitäten im Bereich der Extremismusprävention und der politischen Bildung in Deutschland

Uhl, Andreas; Freiheit, Manuela; Zeibig, Benjamin; Zick, Andreas

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

**Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:**

Hessische Stiftung Friedens- und Konfliktforschung (HSFK)

### Empfohlene Zitierung / Suggested Citation:

Uhl, A., Freiheit, M., Zeibig, B., & Zick, A. (2022). *Evaluationskapazitäten im Bereich der Extremismusprävention und der politischen Bildung in Deutschland*. (PRIF Reports, 9). Frankfurt am Main: Hessische Stiftung Friedens- und Konfliktforschung. <https://doi.org/10.48809/prifrep2209>

### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-ND Lizenz (Namensnennung-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-nd/4.0/deed.de>

### Terms of use:

This document is made available under a CC BY-ND Licence (Attribution-NoDerivatives). For more information see: <https://creativecommons.org/licenses/by-nd/4.0>

# PRIF REPORT

PEACE RESEARCH INSTITUTE FRANKFURT / LEIBNIZ-INSTITUT HESSISCHE STIFTUNG FRIEDENS- UND KONFLIKTFORSCHUNG



ANDREAS UHL // MANUELA FREIHEIT //  
BENJAMIN ZEIBIG // ANDREAS ZICK

**EVALUATIONSKAPAZITÄTEN IM BEREICH  
DER EXTREMISMUSPRÄVENTION UND DER  
POLITISCHEN BILDUNG IN DEUTSCHLAND**

REPORT-REIHE

**PrEval**



PRIF Report 9/2022

# **EVALUATIONSKAPAZITÄTEN IM BEREICH DER EXTREMISMUSPRÄVENTION UND DER POLITISCHEN BILDUNG IN DEUTSCHLAND**

ANDREAS UHL // MANUELA FREIHEIT //  
BENJAMIN ZEIBIG // ANDREAS ZICK

LEIBNIZ-INSTITUT HESSISCHE STIFTUNG FRIEDENS- UND KONFLIKTFORSCHUNG (HSFK)  
PEACE RESEARCH INSTITUTE FRANKFURT (PRIF)

**Coverbild:**

Lizenzinformationen: Joakim Nådell: <https://unsplash.com/photos/K67sBVqLLuw> (Bearbeitung: Anja Feix).

**Textlizenz:**

Creative Commons CC-BY-ND 4.0 (Namensnennung/Keine Bearbeitungen/4.0 International).  
Das Coverbild unterliegt eigenen Lizenzbedingungen.

**Adresse:**

Leibniz-Institut Hessische Stiftung Friedens- und Konfliktforschung (HSFK)  
Baseler Straße 27–31  
60329 Frankfurt am Main  
Telefon: +49 69 95 91 04-0  
E-Mail: [preval@hsfk.de](mailto:preval@hsfk.de)  
<https://www.hsfk.de>

**DOI: 10.48809/prifrep2209**

**ISBN: 978-3-946459-78-1**

## DIE AUTOR:INNEN

*Andreas Uhl* ist Soziologe und seit 2018 wissenschaftlicher Mitarbeiter am *Institut für interdisziplinäre Konflikt- und Gewaltforschung (IKG)* an der Universität Bielefeld. Seine inhaltlichen Arbeitsschwerpunkte liegen im Bereich der Extremismus- und Radikalisierungsprävention.

*Manuela Freiheit* ist Soziologin und seit 2012 wissenschaftliche Mitarbeiterin am *Institut für interdisziplinäre Konflikt- und Gewaltforschung (IKG)* an der Universität Bielefeld. Ihre inhaltlichen Arbeitsschwerpunkte liegen im Bereich der Jugenddelinquenz, der Radikalisierungs- und Konfliktforschung sowie im Bereich der Extremismus- und Radikalisierungsprävention.

*Benjamin Zeibig* ist Politikwissenschaftler und seit 2018 wissenschaftlicher Mitarbeiter am *Institut für interdisziplinäre Konflikt- und Gewaltforschung (IKG)* an der Universität Bielefeld. Seine inhaltlichen Arbeitsschwerpunkte liegen im Bereich der Integrations- und Migrationsforschung, der Friedens- und Konfliktforschung sowie im Bereich der europäischen Außenpolitik im Nahen und Mittleren Osten und Sub-Sahara Afrika.

*Prof. Dr. Andreas Zick* ist seit 2008 Professor für Sozialisation und Konfliktforschung an der Universität Bielefeld. Seit 2013 leitet er als Direktor das *Institut für interdisziplinäre Konflikt- und Gewaltforschung (IKG)*. Das *IKG* erforscht in interdisziplinären Forschungsgruppen die Grundlagen, Phänomene und Wirkungen von innergesellschaftlichen Konflikten und Gewalt. Dazu gehört auch die Prävention von Extremismus und Radikalisierung.

## HSFK-Reportreihe *PrEval*

Über die Evaluation und die wissenschaftliche Begleitung von Maßnahmen der Extremismusprävention und von Angeboten der politischen Bildung wird aktuell in Fachpraxis und Wissenschaft vielfältig diskutiert. Das liegt nicht zuletzt daran, dass über verschiedene Programme auf Ebene des Bundes, der Länder und der Kommunen seit einigen Jahren immer mehr Fördermittel für solche Maßnahmen bereit stehen, die Fachpraxis sich immer vielfältiger, innovativer und professioneller aufstellt und auch die Forschung zu den Phänomenfeldern, aber auch zu Evaluationsmethoden stetig Fortschritte macht. Folgerichtig fordert beispielsweise auch der Kabinettsausschuss zur Bekämpfung von Rechtsextremismus und Rassismus in seinem Maßnahmenpaket eine langfristige Stärkung der Qualität und Wirksamkeit von Präventionsmaßnahmen und knüpft damit an zahlreiche Initiativen von Bundes- und Landesregierungen, von Fachpraxis und Wissenschaft an.

Das Forschungs- und Transfervorhaben *PrEval* (Evaluationsdesigns für Präventionsmaßnahmen – Multimethodische Ansätze zur Wirkungsermittlung und Qualitätssicherung in der Extremismusprävention sowie den Schnittstellen zur Gewaltprävention und politischen Bildung) setzt hier an. Ziel des Projekts ist es, im Dialog aller an Evaluation in diesem Feld beteiligten Akteursgruppen den aktuellen Wissensstand zu Qualitätssicherung und Evaluation in der Extremismusprävention, der Gewaltprävention und der politischen Bildung zu erheben und gemeinsam Evaluationsdesigns, gerade an den Schnittstellen zwischen verschiedenen Präventionsbereichen, Zugängen, Phänomenfeldern und Akteuren, zu erarbeiten. Dabei befasst sich *PrEval* mit Maßnahmen aus der politischen Bildung, der primären, sekundären und tertiären Prävention sowie Schnittstellen zur Gewalt- und Kriminalprävention. *PrEval* nutzt Analyse-, Monitoring- und Mapping-Formate und entwickelt multimethodische Evaluationsdesigns im Rahmen von ausgewählten Pilotstudien. Ein zentraler Mehrwert des Vorhabens ist die Einbindung verschiedener Präventionsakteure – aus der Fachpraxis, den Sicherheitsbehörden, aus der Verwaltung und der Wissenschaft.

Die *PrEval*-Reportreihe fasst die zentralen Ergebnisse des Projekts zusammen. Dabei wird zum einen ein systematischer Überblick über die Evaluationsbedarfe der deutschen Trägerlandschaft in den Präventionsfeldern Islamistischer Extremismus und Rechtsextremismus geboten, die bestehenden Kapazitäten für Evaluation und wissenschaftliche Begleitung von Präventionsmaßnahmen erhoben sowie Evaluationsstrukturen international vergleichend analysiert. Zum anderen werden multimethodische Evaluationsdesigns vorgestellt, die im Rahmen der *PrEval*-Pilotstudien für spezifische Präventionsmaßnahmen in enger Zusammenarbeit mit den Trägern (und teilweise auch Mittelgebern) dieser Maßnahmen entwickelt wurden. Möglichkeiten und Grenzen von wissenschaftlicher Begleitung werden diskutiert. Die zehn Beiträge der Reihe eint ein breites Verständnis von Evaluation, um den Komplexitäten des Felds und der darin verorteten Maßnahmen gerecht zu werden. Gleichwohl setzt jeder Report eigene, dem jeweiligen Thema angepasste Akzente in der Begriffsverwendung.

Das *PrEval*-Projekt wird vom Leibniz-Institut Hessische Stiftung Friedens- und Konfliktforschung (HSFK) koordiniert und von einem Verbund aus Forschungsinstituten und Trägern aus der Fachpraxis gestützt. Mehr Informationen zum Projekt: <https://preval.hsfk.de>

Frankfurt/M. Juli 2022

## Übersicht über alle Titel der *PrEval*-Report-Reihe

**Erfolgsbedingungen und Herausforderungen für die Zusammenarbeit von Fachpraxis und Wissenschaft bei der Evaluation von Angeboten politischer Bildung** von Raphaela Schlicht-Schmälzle // Désirée Theis // Sina Tultschinetski // Melanie Verhovnik-Heinze // Christian Kautz // Maryam Kirchmann. PRIF Report 5/2021.

**Erfahrungen aus der Evaluationsplanung eines Aussteigerprogramms. Voraussetzungen für Wirksamkeitserfassung in der tertiären Extremismusprävention** von Mona Klöckner // Svetla Koynova // Johanna Liebich // Lisa Neef. PRIF Report 6/2021.

**Webvideos und Livestream zur Aufklärung über Desinformation. Evaluationsansätze zu digitalen Formaten in der politischen Bildung** von Désirée Theis // Melanie Verhovnik-Heinze // Stefan Kroll // Keshia Kierek. PRIF Report 8/2021.

**Multimethodische Evaluationsdesigns: eine Erhebung der sozialwissenschaftlichen Praxis** von Mona Klöckner // Melanie Verhovnik-Heinze // Raphaela Schlicht-Schmälzle // Reiko Nakamura // Julius Strunk. PRIF Report 1/2022.

**Klientenzentrierte Evaluation in Multi-Agency-Settings der Extremismusprävention. Möglichkeiten und Grenzen eines wirkungsorientierten Vorgehens** von Susanne Johansson // Julian Junk // Johanna Liebich // Dennis Walkenhorst. PRIF Report 6/2022.

**Monitoring, Evaluation und Lernen: Erfahrungen und Bedarfe der Fachpraxis in der Prävention von Rechtsextremismus und Islamismus** von Svetla Koynova // Alina Mönig // Matthias Quent // Vivienne Ohlenforst. PRIF Report 7/2022.

**Evaluation und Qualitätssicherung in der Kooperation sicherheitsbehördlicher und zivilgesellschaftlicher Akteure in der sekundären und tertiären Extremismusprävention** von Julian Junk // Svetla Koynova // Vivienne Ohlenforst // Maximilian Ruf // Lea Scheu. PRIF Report 8/2022.

**Evaluationskapazitäten im Bereich der Extremismusprävention und der politischen Bildung in Deutschland** von Andreas Uhl // Manuela Freiheit // Benjamin Zeibig // Andreas Zick. PRIF Report 9/2022.

### Weitere geplante Titel in dieser Reihe

**Evaluation der politischen Bildung im Jugendstrafvollzug – Ansätze, Chancen, Herausforderungen** von Sina Tultschinetski // Alisa Rieth // Désirée Theis.

**Developmental Evaluation zur Begleitung von Entwicklungsprozessen in Projekten der Radikalisierungsprävention – ein Anwendungsbeispiel** von Olga Schmidt // Johanna Liebich // Mona Klöckner // Dennis Walkenhorst // Bacem Dziri.

### Außerdem im Rahmen des *PrEval*-Projekts als GPPi-Study erschienen

**Extremismusprävention evaluieren: Institutionelle Strukturen im internationalen Vergleich** von Ase-na Baykal // Sarah Bressan // Julia Friedrich // Giulia Pasquali // Philipp Rotmann // Marie Wagner.



Evaluationen erfüllen vielfältige gesellschaftliche Funktionen und gewinnen auch zunehmend in der Extremismusprävention und ihren Schnittstellenbereichen zur politischen Bildung und Gewaltprävention\* an Bedeutung. Einer Analyse der vorhandenen Evaluationskapazitäten wurde bislang allerdings wenig Aufmerksamkeit geschenkt. Diese Kapazitäten sind jedoch notwendig, um nützliche, durchführbare und genaue Evaluationen realisieren zu können. Auch können sie wichtige Anhaltspunkte dafür liefern, inwieweit sich eine Evaluationskultur bereits ausdifferenziert hat und inwieweit die notwendigen Ressourcen vorhanden sind, um die vielfältigen an Evaluationen gestellten Ansprüche zu erfüllen.

Der vorliegende Report setzt sich daher zum Ziel, die derzeit vorhandenen Evaluationskapazitäten in Deutschland im Bereich der Extremismusprävention mit den Schnittstellen zur politischen Bildung und Gewaltprävention genauer zu bestimmen und aktuelle Entwicklungspotenziale aufzuzeigen. Dazu wurde eine bundesweite Telefonbefragung von 104 Evaluationsakteuren im Bereich der Fremd- (37 Akteure) und Selbstevaluation (67 Akteure) durchgeführt.

Auf der *strukturellen Ebene* zeigt sich, dass die Mittel für Evaluationen vor allem von staatlicher Seite bereitgestellt werden. Diese hat damit einen großen Einfluss auf die Programmsteuerung und -gestaltung und auch auf die Präventionsangebote. Teilweise sind die Mittel und die für Evaluationen zur Verfügung stehende Zeit allerdings knapp bemessen, was deren Handlungsspielraum einschränkt und sich negativ auf die Detailtiefe von Evaluationsergebnissen auswirken kann. Insgesamt sind daher neben finanziellen insbesondere auch mehr zeitliche Ressourcen notwendig, um den hohen Erwartungen gerecht werden zu können.

Eine zu starke Verwaltungssteuerung mit Fokus auf *Outputs* bzw. *Output*-Daten, (z.B. Teilnehmendenzahlen) bindet Evaluationskapazitäten, vor allem dann, wenn diese Daten sehr kleinteilig ermittelt werden müssen. Neben einer zielgerichteten Erhebung sollte auch mehr auf *Outcomes* fokussiert werden, also z.B. auf positive Veränderungen bei den Zielgruppen. Dazu ist es hilfreich, Evaluationen frühzeitig in die Durchführung von Projekten, Maßnahmen und Programmen einzubeziehen, damit sinnvolle und überprüfbare Ziele und Erfolgskriterien formuliert werden können.

Die Förderstrukturen haben aber auch auf der *Organisationsebene* einen starken Einfluss auf die Evaluationskapazitäten und somit auch den Evaluationsmarkt mitgeprägt. Die Kapazitäten haben sich demzufolge auch zu großen Teilen aufgrund des spezifischen Nachfragebedarfs der Förderstrukturen herausgebildet.

Rund ein Drittel fremdevaluierender Evaluationskapazitäten werden dabei von außeruniversitären Forschungseinrichtungen gestellt, gewinnorientierte Unternehmen und universitäre Einrichtungen bilden jeweils etwa ein Viertel der Kapazitäten ab. Vor allem Letztere könnten im Bereich der Evaluationsforschung noch stärker in die Methodenentwicklung eingebunden werden, die Weiter-

---

\* Die politische Bildung bzw. Gewaltprävention sind eigenständige und von der Extremismusprävention abzugrenzende – wenn auch nicht trennscharf voneinander abgrenzbare – Themenfelder. Gleichzeitig können diese aber dennoch einen direkten oder indirekten präventiven Beitrag leisten. Für eine vertiefende Diskussion siehe hierzu z.B. Becker et al. 2020: 165–185.

entwicklung von Ansätzen vorantreiben und vorhandene Forschungslücken schließen. Evaluationsforschung sollte dabei institutionell und konzeptionell enger mit universitärer Forschung verknüpft werden als bisher.

Einrichtungen, die Selbstevaluationen durchführen, müssen aufgrund der doppelten Rollenanforderung von Selbstevaluierenden die Verantwortlichkeiten innerhalb der eigenen Organisation genau klären. Dabei müssen die dafür benötigten personellen und finanziellen Evaluationskapazitäten zur Verfügung gestellt werden, um auch in diesem Bereich den Standards und Gütekriterien von Evaluationen nachkommen zu können.

Die operative Durchführung von Evaluationen erfordert eine große fachliche Expertise auf der *Akteursebene*. Da in Deutschland aktuell nur vereinzelte Ausbildungsmöglichkeiten vorhanden sind, findet die Professionalisierung der Evaluationstätigkeit eher in der Berufspraxis statt. Im Vergleich mit anderen Ländern bestehen hier noch Lücken beim Ausbau von Evaluationskapazitäten. Bei Selbstevaluationen wird meistens auf schon vorhandene Personalressourcen zurückgegriffen, weswegen die Weiterbildung und laufende Qualifizierung einen hohen Stellenwert haben. Befristete Projektlaufzeiten und eine bisher noch nicht auf Dauer ausgelegte Förderstruktur erfordern eine Aufstockung von Kapazitäten.

Vor diesem Hintergrund sollten Selbstevaluationen durch externe Fachkräfte unterstützt werden, um dadurch interne Evaluationsfähigkeiten und -kapazitäten auszubauen. Von solchen – bestenfalls institutionalisierten – Kooperationen würden gleichzeitig auch externe Evaluationsakteure profitieren. Feldkenntnisse auf Seiten der Evaluierenden und Methodenkenntnisse auf Seiten der Evaluierten ließen sich so auf Dauer gegenseitig stärken und dadurch generiertes Wissen könnte in die universitäre Ausbildung und Forschung fließen. Für die Evaluationskapazitäten würde das eine schnellere und wissenschaftlich fundierte Expertisenbildung bedeuten, mit der schneller abgestimmt auf Veränderungen von Radikalisierungsphänomenen reagiert werden könnte.

Die von der Prävention adressierten und von der Evaluation in den Blick genommenen Phänomene treten auch international in Erscheinung und haben dort zu unterschiedlich ausgeprägten Evaluationskapazitäten geführt. Im letzten Abschnitt des Reports wird deshalb eine internationale Verortung anhand einer weiteren im *PrEval*-Verbund durchgeführten Studie vorgenommen. Für Deutschland zeigen sich hier starke Verbesserungspotenziale im Bereich der Fehlerkultur – so haben die Evaluierten zum Teil große Bedenken, bei negativen Evaluationsergebnissen nicht mehr weiter gefördert zu werden. Evaluationsergebnisse sollten außerdem auch für Verbesserungsprozesse in den fördermittelgebenden Institutionen genutzt werden. Generell sollte Evaluation aber noch stärker als strukturbildendes Element der Prävention von Anfang an mitgedacht und in Präventionsbemühungen einbezogen werden. Die entsprechenden Kapazitäten hierfür müssen auch auf struktureller Ebene aufgebaut und weiterentwickelt werden.

Weitere Dimensionen zur Beschreibung und Bestimmung der Evaluationskapazitäten sowie deren Weiterentwicklungspotenzialen finden sich auch auf der *PrEval*-Plattform ([preval-plattform.de](http://preval-plattform.de)).

1. Einleitung	1
2. Begriffe und Konzepte	2
2.1 Evaluationskapazitäten	2
2.2 Interne und externe Evaluation	4
3. Methodisches Vorgehen	6
3.1 Stichprobe	6
3.2 Datenbasis	7
3.3 Erhebung und Feldphase	8
4. Eine Systematisierung der Evaluationskapazitäten in Deutschland	9
4.1 Strukturelle Ebene	10
4.2 Organisationsebene	14
4.3 Akteursebene	22
4.4 Die <i>PrEval</i> -Plattform – ein Datenvisualisierungstool für die Evaluation	28
5. Internationale Verortung	32
6. Fazit und Ausblick	36
Literatur	38



## 1. EINLEITUNG

Bei der Prävention und Bekämpfung von Extremismus und Gewalt hat die Evaluation von Maßnahmen, Projekten und Programmen eine zentrale Funktion. Auf Basis empirisch gewonnener Daten kann sie wichtige Erkenntnisse zur Weiterentwicklung und Steuerung der Präventions- wie Interventionspraxis liefern. Zugleich stellt sie aber auch ein wichtiges Bindeglied zwischen den an der Prävention und Intervention beteiligten Akteuren aus Praxis, Politik und Wissenschaft dar.

Während im Mittelpunkt aktueller Diskussionen insbesondere die Vor- und Nachteile verschiedener Evaluationsansätze sowie deren Aussagekraft und Gegenstandsangemessenheit stehen, wurde den vorhandenen Kapazitäten sowie der Frage, welche Kapazitäten notwendig sind, um eine effektive und konstruktive Evaluationspraxis im Bereich der Extremismusprävention und Demokratieförderung<sup>1</sup> zu ermöglichen, bislang wenig Aufmerksamkeit zuteil. Zwar sind in den großen Bundesprogrammen – etwa *Demokratie leben!* oder *Zusammenhalt durch Teilhabe* – Evaluationen in den Programmen direkt angelegt und auch sonst scheint Evaluation bereits im Alltag der Praxisarbeit weit verbreitet und etabliert zu sein (Koynova et al. 2022). Gleichzeitig liegt bis dato jedoch wenig Wissen zu bestehenden Evaluationskapazitäten in Deutschland vor. Auch fehlte es bisher an einer systematischen Übersicht über die Evaluationskapazitäten in diesem Bereich. Das überrascht umso mehr, da die Qualität von Evaluationen stark von den ihr zur Verfügung stehenden Kapazitäten abhängt. Anders formuliert: Eine Bestimmung des Evaluationsgeschehens von Präventionsmaßnahmen ohne eine Analyse von vorhandenen, fehlenden, erwünschten wie erwarteten Kapazitäten ist problematisch. Denn nur mithilfe einer detaillierten Übersicht lässt sich ermitteln, inwiefern die vielfältigen Erwartungen, die an die Qualitätssicherung allgemein und die Evaluationspraxis im Besonderen gestellt werden, mit den gegenwärtigen Kapazitäten überhaupt erfüllt werden können.

Ziel des im Rahmen des *PrEval*-Forschungsverbunds durchgeführten Teilprojekts *Mapping von Evaluationskapazitäten* war es daher, die derzeit vorhandenen Evaluationskapazitäten im Bereich der Extremismusprävention sowie den Schnittstellen zur Gewaltprävention und politischen Bildung in Deutschland zu erfassen und diese in eine öffentlich zugängliche digitale Plattform zu überführen. Dazu hat das *Institut für interdisziplinäre Konflikt- und Gewaltforschung (IKG)* an der Universität Bielefeld eine bundesweite Befragung von Evaluationsakteuren durchgeführt und neben der systematischen Bestandsaufnahme der Evaluationskapazitäten auch die unterschiedlichen Ansätze und Methoden der Evaluierenden in den Blick genommen.

Bevor die Ergebnisse dieser systematischen Bestandsaufnahme diskutiert werden, sollen zunächst jedoch die bereits eingeführten Begriffe und Konzepte genauer definiert, das methodische Vorgehen sowie die der Befragung zugrundeliegende Datenbasis vorgestellt werden. In der detaillierten Analyse der Evaluationskapazitäten in Abschnitt 4 dieses Reports sollen dabei insbesondere die strukturelle Ebene, die Organisationsebene und individuelle Akteursebene genauer betrachtet sowie die zentralen Ergebnisse berichtet werden. Alle weiteren Kategorien der bundesweiten Befragung, die im Report keinen Platz gefunden haben, sind über die digitale *PrEval*-Plattform öffentlich ein-

---

1 Dies schließt auch die Schnittstellen zur politischen Bildung und Gewaltprävention ein.

sehbar.<sup>2</sup> Dort lassen sich die unterschiedlichen Kategorien mithilfe von verschiedenen Analyse- und Visualisierungswerkzeugen in Beziehung setzen und anhand eigener Fragestellungen tiefergehend analysieren. Abschließend werden die dargestellten Befunde in den Kontext der ebenfalls in *PrEval* durchgeführten internationalen Studie des *Global Public Policy Institute (GPPi)* gesetzt und mit den Praktiken und Strukturen anderer Länder verglichen (Baykal et al. 2021).

## 2. BEGRIFFE UND KONZEPTE

Zunächst sei festgehalten, dass sich der *PrEval*-Forschungsverbund mit einem weiten Phänomenbereich beschäftigt. Dieser umfasst politische wie religiöse Extremismen, adressiert Maßnahmen der universellen, selektiven und indizierten Prävention und schließt dabei auch die Schnittstellen zur politischen Bildung und Gewaltprävention ein. Obwohl die Meinungen zur Abgrenzung von politischer Bildung und Prävention auseinandergehen (siehe hierzu z.B. Becker et al. 2020: 165–185), so besteht doch Einvernehmen darüber, dass in diesen Schnittstellenbereichen präventive Wirkungen erzielt werden können – genauso, wie in Präventionsangeboten erlernte Inhalte wiederum „Momente langfristiger Bildungsprozesse sein können“ (Milbradt et al. 2019: 157). Dieser in *PrEval* untersuchte weite Phänomenbereich spiegelt sich auch im Mapping der Evaluationskapazitäten wider. Die Bestimmung von Evaluationskapazitäten erfolgt dementsprechend nicht phänomenspezifisch, auch wenn spezifische Extremismusformen eventuell im Detail unterschiedliche Evaluationsbedingungen nach sich ziehen.

### 2.1 EVALUATIONSKAPAZITÄTEN

Um bestimmen zu können, welche Evaluationskapazitäten im Bereich der Extremismusprävention und Demokratieförderung in Deutschland zur Verfügung stehen, muss zunächst das Konzept der *Kapazitäten* näher bestimmt werden. Zumeist werden Evaluationskapazitäten in diesem Kontext eher indirekt in der Literatur umrissen. Gut ausgearbeitete Konzepte finden sich vor allem im Bereich der Entwicklungszusammenarbeit. Dort hat sich in den letzten Jahren eine starke Evaluationskultur ausgebildet und es konnten bereits vielfältige Professionalisierungsprozesse angeregt werden (Stockmann 2018). Ein Teil dieser Professionalisierung betrifft den Auf- und Ausbau von Evaluationskapazitäten (Stockmann 2018: 1–3). Unter einem solchen „Kapazitätenaufbau“ werden im Bereich der Entwicklungszusammenarbeit meist Prozesse zur Etablierung von Qualitätssicherungsprozessen gefasst. Über diese Prozesse werden Fähigkeiten zur Durchführung und Nutzung von Evaluationen freigesetzt, gestärkt, angepasst und in Routinen überführt. Damit können Rechenschaftspflichten erfüllt, aber auch Lernprozesse wirksam unterstützt werden (vgl. Stockdill 2002, OECD 2021).

Unter Berücksichtigung der Forschungsliteratur lassen sich bei diesem sogenannten *Evaluation Capacity Development* bzw. *Building* verschiedene Gemeinsamkeiten der jeweiligen Ansätze herausar-

---

<sup>2</sup> <http://preval-plattform.de> (letzter Zugriff: 17.05.2022).

beiten und die wichtigsten Kriterien für die Vermessung von Evaluationskapazitäten im Bereich der Extremismusprävention sowie den Schnittstellen zur Gewaltprävention und politischen Bildung ableiten.

Den verschiedenen Ansätzen ist dabei gemein, dass Evaluationskapazitäten jeweils auf verschiedenen *Ebenen* ausgebildet bzw. weiterentwickelt werden können. Nach einer gängigen Definition von Stockdill und Kollegen umfasst dies „die gezielte Arbeit zur kontinuierlichen Schaffung und Aufrechterhaltung umfassender organisatorischer Prozesse, welche qualitativ hochwertige Evaluationen und ihre Anwendung zur Routine werden lassen“ (Stockdill et al. 2002: 14; Übersetzung des Verfassers). Diese Definition fokussiert vor allem die Organisationsebene. Zu dieser Ebene gehören beispielsweise vorhandene Personalressourcen oder auch die finanziellen Ressourcen innerhalb einer Organisation (vgl. hierzu auch Labin et al. 2012: 11). Andere Ansätze nehmen die individuelle Ebene von Evaluationskapazitäten in den Blick und beziehen sich dabei vor allem auf das Wissen und die Fähigkeiten einzelner Personen, also der Mitarbeiter:innen innerhalb einer Organisation (Labin et al. 2012). Bei Stockmann werden diese beiden Ebenen noch um eine systemische Ebene ergänzt: „Diese, eine Organisation umgebenden Systeme, können z.B. politische, ökonomische, soziale, kulturelle, rechtliche, internationale oder andere Subsysteme sein“ (Stockmann 2018: 10). Solche Systeme können beeinflussen, welche Evaluationskultur gelebt wird oder welche institutionellen Evaluationsstrukturen bestehen.

Angelehnt an dieses Verständnis versteht der vorliegende Bericht die Evaluationskapazitäten im Bereich der Extremismusprävention und Demokratieförderung als multiperspektivisch und auf unterschiedlichen Ebenen angesiedelt, d.h. als selbst- und fremdwahrgenommene Kapazitäten auf der Ebene von Akteuren, Organisationen und Strukturen. Auch lassen sich vor dem Hintergrund dieses Verständnisses für die anschließende Analyse und Systematisierung der Evaluationskapazitäten in Deutschland insbesondere die in Abbildung 1 dargestellten Fragen und Indikatoren ableiten. Dabei wird nicht zuletzt dem Umstand Rechnung getragen, dass Evaluation in der Extremismusprävention ein noch relativ junges Themenfeld ist und bislang nicht auf eigene oder gar zertifizierte Kriterien zurückgegriffen werden kann.



Abb. 1: Die unterschiedlichen Ebenen der Evaluationskapazitäten.

Bevor die Systematisierung der Evaluationskapazitäten in Kapitel 4 entlang dieser Ebenen und Fragen erfolgt, soll über die verschiedenen Ebenen der Evaluationskapazitäten hinaus zwischen internen und externen Evaluationsperspektiven unterschieden werden. Diese stellen zwei wichtige Bereiche der Evaluationskapazitäten dar.

## 2.2 INTERNE UND EXTERNE EVALUATION

Bei der Durchführung von Evaluationen lassen sich generell zwei Evaluationsperspektiven einnehmen: Der evaluative Blick kann entweder von außerhalb der zu evaluierenden Organisation (*externe Evaluation*) oder von innerhalb erfolgen (*interne Evaluation*).

*Externe Evaluationen* werden dabei von unabhängigen Evaluator:innen außerhalb der Organisation durchgeführt, deren Projekte, Maßnahmen oder Programme evaluiert werden sollen.

*Interne Evaluationen* werden hingegen von Personen innerhalb der eigenen Organisation durchgeführt. Hierbei werden zwei Formen unterschieden:

- Eine *interne Selbstevaluation* wird von Personen umgesetzt, die auch an der Projekt- bzw. Maßnahmenkonzeption und/oder ihrer Durchführung beteiligt waren bzw. sind.
- Eine *interne Fremdevaluation* wird zwar ebenfalls von Personen innerhalb der eigenen Organisation durchgeführt, diese sind aber außerhalb des zu evaluierenden Projekts bzw. der zu evaluierenden Maßnahme tätig.

Interne und externe Evaluationen können sich dabei in vielerlei Hinsicht ergänzen. Auch können beide Perspektiven zur Verfestigung einer Evaluationskultur beitragen und spielen in Qualitätssicherungsprozessen eine wichtige Rolle (vgl. Koynova et al. 2022). Es ist jedoch darauf zu achten, dass sowohl externe als auch interne Evaluationen nach professionellen Standards durchgeführt werden. Die *Deutsche Gesellschaft für Evaluation (DeGEval)* gibt hierzu für beide Bereiche Standards aus (DeGEval 2004; DeGEval 2017), die auch für die Extremismusprävention und deren Evaluation zunehmend an Relevanz gewinnen.

Laut dieser Standards sollen Evaluationen vier grundlegende Eigenschaften erfüllen: Sie sollen nützlich, durchführbar, fair und genau sein (DeGEval 2004; DeGEval 2017). Daran schließen sich jeweils mehrere Einzelstandards an, welche die grundlegenden Kriterien noch einmal konkretisieren. Unterschiedliche Anforderungen an die Evaluierenden – etwa hinsichtlich Kompetenz und Glaubwürdigkeit – sind hier genauso Bestandteil wie ein angemessenes Verfahren, eine angemessene Analyse qualitativer und quantitativer Informationen, aber auch die Offenlegung von Ergebnissen und Berichten (DeGEval 2004; DeGEval 2017).

Aufgrund der doppelten Rollenanforderung bei Selbstevaluationen, bei denen Praxisakteure einerseits als Akteure innerhalb ihres Arbeitsfeldes sowie andererseits als Beobachtende und Bewertende auftreten (DeGEval 2004: 7), kann eine Unparteilichkeit wie im Bereich der externen Evaluation allerdings nicht erwartet werden. Im Großen und Ganzen decken sich die Standards der Selbstevaluation

jedoch mit den allgemeinen bzw. den Standards für Fremd- bzw. externe Evaluationen. Selbstevaluationen können aber auch Vorteile mit sich bringen, so etwa

„die Vertrautheit mit den Interna des Arbeitsfeldes, die unmittelbare Integrationsmöglichkeit der Ergebnisse in den Alltag, die Qualifizierung der MitarbeiterInnen sowie die Spannung und Neugier auf Antworten für selbst gestellte Fragen“ (Müller-Kohlenberg 2006: 75).

Gleichzeitig besteht bei Selbstevaluationen aber die Gefahr, dass Eigeninteressen dazu verleiten, die selbstevaluierte Arbeit im Sinne des Erwünschten zu präsentieren (Müller-Kohlenberg 2006: 76). Deshalb ist die Selbstevaluation auch nicht für alle Anwendungsfälle gleichermaßen geeignet, z.B., wenn es darum geht, eine Weiterführungsentscheidung zu einem Projekt aus Evaluationsergebnissen abzuleiten. In solchen Fällen würde sich eine Fremdevaluation besser eignen. Selbstevaluationen können aber als Mittel zur kontinuierlichen Verbesserung der eigenen Arbeit und als Teil des Qualitätsmanagements Fremdevaluationen sinnvoll ergänzen.

Externe Evaluationen und Selbstevaluationen folgen demnach ihren jeweils eigenen Handlungslogiken und sollten deshalb auch unabhängig voneinander betrachtet werden. Nach Stockmann

„werden externe Evaluationskapazitäten eher für die unabhängige Analyse im Dienste der gesellschaftlichen Aufklärung und der demokratischen Legitimierung genutzt, während sich für die Programmsteuerung sowohl interne als auch externe Evaluationskapazitäten einsetzen lassen“ (Stockmann 2008: 5).

Je unabhängiger Evaluationen durchgeführt werden können, desto „glaubwürdiger ist ihr Beitrag zur gesellschaftlichen Aufklärung, zur demokratischen Legitimitätsbeschaffung und zur Programmsteuerung“ (Stockmann 2008: 6). Dies gilt insbesondere für externe Evaluationen durch professionell geschultes Personal.

Im Weiteren werden vor allem die Begriffe *Fremdevaluation* (synonym mit externer Evaluation) und *Selbstevaluation* verwendet. Die interne Fremdevaluation wird dabei ausgeblendet. Damit wird ein Ergebnis der bundesweiten Befragung bereits vorweggenommen: In den Interviews hat sich gezeigt, dass sich bei internen Evaluationen die Akteurs- und Evaluationsebenen nicht eindeutig abgrenzen ließen und die interviewten Einrichtungen daher insgesamt im Bereich der Selbstevaluation zu verorten waren. Ausschlaggebend sind hierbei oft personelle Gründe – viele Angestellte arbeiten in unterschiedlichen Rollen und projektübergreifend, sodass sich eine „echte“ interne Fremdevaluation zum Teil nur schwer umsetzen lässt.

### 3. METHODISCHES VORGEHEN

Um die Evaluationskapazitäten im Bereich der Extremismusprävention sowie den Schnittstellen zur Gewaltprävention und politischen Bildung in Deutschland zu erheben, wurde eine bundesweite telefonische Befragung von Evaluationsakteuren durchgeführt. Im ersten Schritt wurden dabei die Evaluationsakteure entlang von Bundes- und Landesprogrammen bis hinab auf die kommunale Ebene identifiziert und erfasst. Hierzu wurden einerseits ausführliche On- und Offlinerecherchen durchgeführt, aber auch wichtige Gatekeeper (wie z.B. Landes-Demokratiezentren) und andere Stakeholder und Quellen in die Identifizierungsphase einbezogen. Vervollständigt wurde der Datenpool zudem durch Rückmeldungen aus den Bedarfserhebungen mit Praxisakteuren der Extremismusprävention, die im Rahmen von *PrEval* in Zusammenarbeit mit dem *Kompetenznetzwerk Islamistischer Extremismus (KN:IX)* und dem *Kompetenznetzwerk Rechtsextremismusprävention (KompRex)* durchgeführt wurden (Koynova et al. 2022). Zusätzlich konnten mithilfe der im Rahmen des *MAPEX*-Forschungsprojekts<sup>3</sup> gewonnenen Daten weitere selbstevaluierende Präventionsakteure identifiziert und in die Erhebung einbezogen werden (siehe hierzu Freiheit et al. 2021a: 49–78 bzw. [www.mapex-projekt.de](http://www.mapex-projekt.de)).

#### 3.1 STICHPROBE

Hinsichtlich der Auswahlkriterien für die zu befragenden Einrichtungen bzw. Akteure ergaben sich folgende Unterschiede in Bezug auf die Fremd- und Selbstevaluation:

##### FREMDEVALUATION

- Hier wurde eine möglichst vollständige Erhebung fremdevaluierender Einrichtungen im Bereich der Extremismusprävention sowie den Schnittstellen zur Gewaltprävention und politischen Bildung in Deutschland angestrebt.
- Nicht in den Blick genommen wurden selbstständig tätige Einzelpersonen bzw. Freiberufler:innen oder Beratungsunternehmen, die eher im Sinn einer Unternehmensberatung Geschäftsprozesse o.ä. evaluieren oder in diesem Rahmen beraten.

Insgesamt konnten 90 Einrichtungen identifiziert werden, die zunächst den oben genannten Aufnahmekriterien entsprachen.

##### SELBSTEVALUATION

Ausgewählt wurden:

- Träger, die im Rahmen der *PrEval*-Bedarfserhebungen rückgemeldet wurden;

---

3 *MAPEX – Mapping und Analyse von Präventions- und Distanzierungsprojekten im Umgang mit islamistischer Radikalisierung* war ein Forschungs-Verbundprojekt (12/2017–02/2021), das zum Ziel hatte, einen bundesweiten Überblick sowohl über sämtliche öffentlich zugängliche Maßnahmen der Prävention als auch der Deradikalisierung und Distanzierung im Bereich des islamistischen Extremismus zu entwickeln. In dessen Rahmen fand eine bundesweite telefonische Befragung von Präventionsakteuren statt. Hierbei wurde unter anderem gefragt, ob bereits Selbstevaluationen für die realisierten Projekte oder Maßnahmen durchgeführt wurden. Weitere Informationen finden sich unter <https://www.mapex-projekt.de/projektinformationen> (letzter Zugriff: 16.05.2022).

- im *MAPEX*-Datensatz enthaltene Träger,
  - die angegeben hatten, bereits eine Selbstevaluation durchgeführt zu haben und in mindestens einem der Phänomenbereiche des islamistischen Extremismus, Rechtsextremismus oder Linksextremismus tätig waren;
  - deren Angebote unterschiedliche Reichweiten aufweisen (von online, deutschlandweit, bundeslandweit, kommunal bis hin zu in einer Stadt/in einem Stadtteil tätigen Projekten bzw. Maßnahmen), um eine möglichst breite Abdeckung zu erreichen.
- Einrichtungen, die im Interview angaben, nur interne Fremdevaluationen durchzuführen, wurden der Selbstevaluation zugerechnet, da sich insbesondere hier die Akteurs- und Evaluationsebene nicht eindeutig abgrenzen ließen und diese eher als Selbstevaluationsakteure zu sehen waren (siehe hierzu auch DeGEval 2004: 6 und Abschnitt „Interne und externe Evaluation“).
- Nicht berücksichtigt wurden Einzelprojekte oder -maßnahmen, außer, wenn diese in Form einer Fach- oder Beratungsstelle auftraten.

Auch konnte aufgrund der kurzen Laufzeit des *PrEval*-Projekts im Bereich der Selbstevaluation bzw. internen Evaluation keine Zufallsstichprobe über die gesamte Trägerlandschaft in Deutschland gezogen werden, sodass hier auf die drei oben genannten internen Datensätze zurückgegriffen wurde. Bei den so ausgewählten Trägern war dann jeweils bekannt, dass diese bereits eine interne Evaluation bzw. Selbstevaluation durchgeführt hatten. Insgesamt wurden 183 Einrichtungen ausgewählt.

### 3.2 DATENBASIS

Die insgesamt 273 so ausgewählten und identifizierten Einrichtungen der Fremd- und Selbstevaluation wurden im zweiten Schritt meist auf Geschäftsführungsebene telefonisch kontaktiert, um zu validieren, ob Fremd- bzw. Selbstevaluationen im Bereich der Extremismusprävention und/oder der politischen Bildung bzw. Gewaltprävention auch tatsächlich durchgeführt wurden. Von Seiten der Interviewer:innen wurden dabei keine Definitionen vorgegeben, was unter einer Evaluation zu verstehen sei, sondern es wurde das jeweilige Selbstverständnis erfragt.

Insgesamt gaben hierbei 21% ( $n = 38$ ) der 183 als selbstevaluierend identifizierten Einrichtungen an, dass sie keine Evaluationen im eigentlichen Sinn durchführen, sondern dass im Rahmen der innerhalb der eigenen Organisation durchgeführten Projekte und Maßnahmen nur eine reduzierte Kennzahlenerhebung durchgeführt wird (z.B. Teilnehmer:innenzahlen bei Veranstaltungen) oder nur sehr einfache Feedbackbögen nach den Veranstaltungen ausgegeben werden. Weitere 7% ( $n = 13$ ) führten keine Evaluationen in den im Rahmen von *PrEval* untersuchten Bereichen durch. In 2% ( $n = 4$ ) der Einrichtungen wurden (bisher) keine Evaluationen umgesetzt.

Einige Einrichtungen der Selbstevaluation, die in den Bedarfserhebungen bzw. im *MAPEX*-Datensatz angegeben hatten, Selbstevaluationen durchzuführen, verneinten dies dann aber in den Anbahnungsgesprächen im Rahmen der Kapazitätenerhebung. Dieser Umstand weist auf ein uneinheitliches, sehr breit gefächertes Verständnis von Evaluation hin. Künftige Befragungen sollten diesbezüglich angepasst werden, vor allem standardisierte Ja/Nein-Fragen vermeiden bzw. das jeweilige Selbstverständnis der Einrichtungen berücksichtigen und abfragen.

Im Bereich der Fremdevaluation gaben 13% ( $n = 12$ ) der insgesamt 90 identifizierten Einrichtungen an, nicht über Expertisen in den von *PrEval* fokussierten Bereichen zu verfügen bzw. keine Evaluationen in diesen Bereichen explizit anzubieten.

Zum Teil wurde die Kontaktaufnahme durch die während des Erhebungszeitraums anhaltende Corona-Pandemie erschwert und es ergaben sich vor allem auf Seiten der Träger Zeit- und Ressourcenprobleme, sodass eine Teilnahme nicht immer möglich war.

Von den insgesamt 273 erfassten Einrichtungen konnten jedoch

- 104 Einrichtungen (38%) für ein Interview gewonnen werden, davon 37 (36%) im Bereich der Fremdevaluation und 67 (64%) im Bereich der Selbstevaluation.
- Bei 77 Einrichtungen (28%) stellte sich im Anbahnungsgespräch heraus, dass keine (Selbst-) Evaluationen im eigentlichen Sinn, keine Evaluationen in den von *PrEval* untersuchten Bereichen oder bisher überhaupt keine Evaluationen durchgeführt wurden.
- Bei 48 Einrichtungen (18%) kam kein Kontakt zustande, konnte ein bereits vereinbarter Interviewtermin nicht mehr realisiert werden, gab es die Einrichtung nicht mehr oder es verhinderten sonstige Gründe das Zustandekommen eines Interviews.
- Bei 44 Einrichtungen (16%) war eine Interviewteilnahme aus Zeit- bzw. Ressourcenmangel nicht möglich bzw. diese wurde explizit abgelehnt.

### 3.3 ERHEBUNG UND FELDPHASE

Die Erhebung der Daten erstreckte sich von September 2020 bis Juni 2021 mit einigen nachgelagerten Interviews, sofern im Haupterhebungszeitraum kein Interviewtermin gefunden werden konnte. Der Erhebungsphase ging ein Standard-Pretest voraus, um den Fragebogen auf Verständlichkeit und statistische Eignung hin zu überprüfen und ihn im Anschluss daran für die Haupterhebung zu überarbeiten. In diesem Schritt wurden lediglich Fragenformulierungen präzisiert und Hinweistexte ergänzt, weitere Anpassungen am Fragebogen wurden nicht vorgenommen.

Die Interviews wurden jeweils von den Projektmitarbeiter:innen des *IKG* durchgeführt. Im Schnitt waren fünf Kontaktversuche erforderlich, bevor ein Interview durchgeführt werden konnte. Die durchschnittliche Interviewdauer betrug 59 Minuten, mit einem Maximum von 166 und einem Minimum von 25 Minuten. Die lange durchschnittliche Interviewdauer unterstreicht auch noch einmal den Gesprächs- und Austauschbedarf, der sowohl von den Trägern als auch den evaluierenden Einrichtungen ausging.

Zu den Schwerpunkten des Fragebogens zählte die Erhebung

- der Evaluationsakteure,
- der Auftraggeber,
- der angewandten Evaluationsmethoden, -ansätze und -formen,
- der Qualifikationen der Evaluierenden,

- der Erfolgskriterien und Ziele von Evaluationen,
- der evaluierten Arbeitsbereiche und Präventionsebenen,
- von Herausforderungen, Evaluations- und Forschungsbedarfen aus Sicht der Evaluierenden,
- von Verbesserungspotenzialen bei der Zusammenarbeit mit Mittelgebern bzw. evaluierten Einrichtungen und Projekten,
- aber auch die Erhebung der positiven Entwicklungen der letzten Jahre im Bereich der Evaluation.

Insgesamt wurden 30 verschiedene Kategorien abgefragt.

Der Fragebogen wurde mit der OpenSource-Umfrage-Software *LimeSurvey* erstellt. Dadurch konnten die Antworten der Befragten während des Telefoninterviews direkt über eine grafische Benutzeroberfläche eingegeben und in der Umfragedatenbank gespeichert werden. Die Umfragedatenbank lieferte anschließend die Grundlage für die digitale *PrEval*-Plattform, mit der die Forschungsdaten visualisiert werden.

Diese öffentlich zugängliche Plattform bildet dabei einerseits die Evaluationsakteure auf einer interaktiven Landkarte ab, andererseits stellt sie aber auch Analysemöglichkeiten innerhalb der Plattform bereit, um die Evaluationskapazitäten anhand eigener Fragestellungen analysieren zu können (siehe hierzu Kapitel 4).

#### 4. EINE SYSTEMATISIERUNG DER EVALUATIONSKAPAZITÄTEN IN DEUTSCHLAND

Die Systematisierung der Evaluationskapazitäten soll, wie im Abschnitt *Evaluationskapazitäten* beschrieben, über drei Ebenen erfolgen – die *strukturelle Ebene*, die *Organisationsebene* und die *Akteursebene*. Diesen Ebenen lassen sich verschiedene Kategorien aus dem Fragenkatalog der bundesweiten Befragung zuordnen. Daraus ergibt sich ein Analyseschema, anhand dessen die Evaluationskapazitäten in den nachfolgenden Abschnitten genauer beschrieben werden (Tabelle 1).

ANALYSEEBENE	KATEGORIEN ZUR BESCHREIBUNG DER EVALUATIONSKAPAZITÄTEN
Strukturelle Ebene	Finanzierung und Auftraggeber Verwertung von Evaluationsergebnissen
Organisationsebene	Art der evaluierenden Einrichtung Evaluationsgegenstände Evaluationsformen lokale Verortung
Akteursebene	Qualifikationen Evaluationsansätze und -modelle

Tab. 1: Systematisierung der Evaluationskapazitäten.

Zudem soll in den nächsten Abschnitten die Perspektive der fremd- und selbstevaluierenden Einrichtungen an geeigneter Stelle um die Bedarfe und Evaluationserfahrungen von Praxisakteuren der Extremismusprävention ergänzt werden.

In Kooperation mit *KN:IX* und *KompRex* entstanden im Rahmen von *PrEval* – neben der Erhebung der Evaluationskapazitäten – zwei Bedarfserhebungen, welche die Praxisperspektive genauer beleuchten. Hierzu wurden insgesamt 429 Praktiker:innen der Phänomenfelder islamistischer Extremismus ( $n = 106$ ) und Rechtsextremismus ( $n = 323$ ) zu ihren Evaluationserfahrungen, -bedarfen und -anforderungen befragt.<sup>4</sup>

Durch den Vergleich der unterschiedlichen Perspektiven von Evaluations- und Präventionsakteuren lässt sich feststellen, an welchen Stellen Bedarfe und Kapazitäten bereits eine gute Passgenauigkeit entwickelt haben oder die Perspektiven ggf. auch auseinandergehen.

## 4.1 STRUKTURELLE EBENE

### 4.1.1 FINANZIERUNG UND AUFTRAGGEBER

Um Evaluationen sinnvoll durchführen zu können, bedarf es auf der strukturellen Ebene für die evaluierenden Einrichtungen eines Mindestmaßes an finanzieller Sicherheit.

Wie die Ergebnisse der bundesweiten Erhebung zeigen, werden die Mittel dabei zum überwiegenden Teil von staatlicher Seite bereitgestellt. So gaben 76% ( $n = 51$ ) der selbstevaluierenden Einrichtungen an, staatliche Mittel für intern durchgeführte Evaluationen erhalten zu haben. Im Bereich der Fremdevaluation lag der Anteil sogar bei 95% ( $n = 35$ ). Dadurch, dass staatliche Einrichtungen als Mittelgeber und damit auch gleichzeitig als Auftraggeber auftreten, sofern Evaluationsanteile für die Zuwendung verpflichtend sind, obliegt ihnen zumeist auch die Ausgestaltung des Leistungsumfangs der Fremd- bzw. Selbstevaluationen.

Auch betonten mehr als die Hälfte aller befragten Einrichtungen, dass nur zum Teil oder nicht ausreichende Mittel<sup>5</sup> zur Verfügung stehen, um den großen Erwartungshaltungen der Mittelgeber gerecht werden zu können. Die „Zielerwartung an Evaluation [sei] deutlich über den mit dem Budget möglichen Methoden“, so ein Interviewpartner im Bereich der Fremdevaluation. Auch erfolge die Auftragsvergabe primär auf Preisbasis. Laut der befragten fremdevaluierenden Einrichtungen sollte hier jedoch eher der Ideenwettbewerb der Anbieter und nicht der Preiswettbewerb im Vordergrund stehen, sodass sich die vorhandenen Evaluationskapazitäten gegenstandsangemessen einsetzen

---

4 Die Ergebnisse der Bedarfserhebungen werden im PRIF Report „Monitoring, Evaluation und Lernen: Erfahrungen und Bedarfe der Fachpraxis in der Prävention von Rechtsextremismus und Islamismus“ (Koynova et al. 2022) detailliert vorgestellt.

5 Die Antwortskala umfasste die Antwortmöglichkeiten „angemessen“, „eher angemessen“, „teils, teils“, „eher nicht angemessen“, „nicht angemessen“, „weiß nicht“ sowie „keine Angabe“.

ließen. Zudem sei in den Zuwendungen für Präventionsangebote zum Teil nur ein geringer Anteil der Mittel für Evaluationen vorgesehen, dieser sei fix und nicht verhandelbar, was sich zusätzlich einschränkend auf die Ausgestaltung von Evaluationen auswirken würde.

Aber nicht nur die zur Verfügung stehenden finanziellen Mittel, sondern auch der Zeitfaktor spielt nach Aussagen der interviewten Einrichtungen der Fremd- und Selbstevaluation eine wichtige Rolle. So müsse seitens der Mittelgeber mehr Zeit für eine konkrete Bewertung eingeräumt werden, um professionellere und tiefergehende Aussagen über die Ergebnisse einer Evaluation treffen zu können. Aus Sicht der Befragten sollten die Projektlaufzeiten auch im besten Fall um eine nachgelagerte Reflexionsphase verlängert werden bzw. diese mitdenken. Oft würde die Projektarbeit zeitgleich mit der Zielgruppenarbeit enden. Hier zeige sich immer noch ein deutlicher Nachbesserungsbedarf bei der Programmgestaltung und -finanzierung.

Zudem sei Evaluation noch viel zu selten bereits in der Programmkonzeption angelegt. In den Interviews wurde in diesem Zusammenhang auch der zu große Einfluss der Verwaltungssteuerung genannt, welche eher zu einer *Output*-Steuerung von Programmen führt und weshalb sehr stark auf Kennzahlen geachtet werden würde. Akteure der Fremdevaluation empfehlen hier einen Perspektivwechsel, bei dem weniger von den Maßnahmen her gedacht und sich mehr an Wirkungen orientiert werden sollte.

„Wenn ich viele Veranstaltung mit tausenden Teilnehmenden durchgeführt habe, [...] heißt das noch lange nicht, dass die Maßnahme erfolgreich war. [...] Leider ist das, was ich nach wie vor stark sehe, dass es um solche messbaren *Outputs* geht und weniger um die Qualität, die damit verbunden ist“,

so ein Interviewpartner der Fremdevaluation. Wiederholt wurde auch darauf hingewiesen, dass Evaluationskapazitäten entlastet werden könnten, indem bei Evaluationen nur dort *Output*-Daten erhoben werden, wo diese für die Beurteilung auch erforderlich sind. Laut der Befragten sollten diese auch nicht als alleinige Grundlage für weitreichende Förderentscheidungen und damit für die Programmsteuerung herangezogen werden.

#### **4.1.2 VERWERTUNG VON EVALUATIONSERGEBNISSEN**

Wie in Kapitel 2 beschrieben, ist für den Aufbau und die Stärkung von Evaluationskapazitäten weiterhin entscheidend, dass Evaluation keinen Selbstzweck verfolgen. Sie müssten vielmehr in routinierteres Handeln überführt werden, um Lernprozesse bestmöglich unterstützen zu können. Wesentlich in diesem Zusammenhang ist die aktive Nutzung von Evaluationsergebnissen, um Projekte, Maßnahmen und Programme stetig zu verbessern und solche Lernprozesse innerhalb der evaluierten Einrichtungen anzuregen.

Gefragt nach der Nutzung von Ergebnissen und Erkenntnissen von Evaluationen, gaben 93% ( $n = 97$ ) der interviewten Einrichtungen der Fremd- und Selbstevaluation an,<sup>6</sup> dass diese zur Weiterentwicklung von Methoden und Maßnahmen eingesetzt werden. Fast genauso so häufig wurde mit 92% ( $n = 96$ ) die Optimierung der Konzeptqualität von Projekten und Programmen genannt. Für Steuerungszwecke von Projekten und Programmen werden Evaluationsergebnisse von 79% ( $n = 82$ ) der befragten Einrichtungen genutzt.

Insgesamt unterscheidet sich die Verwendung von Evaluationsergebnissen für die Bereiche der Fremd- und Selbstevaluation nicht stark voneinander. Lediglich die Konzeption neuer Maßnahmen und das Treffen strategischer Entscheidungen über Kooperationspartner:innen wird von selbstevaluierenden Einrichtungen häufiger als von fremdevaluierenden Einrichtungen ausgewiesen, was sich durch die jeweilige Evaluationsperspektive erklären lässt. So sind fremdevaluierende Einrichtungen nicht immer in die der Evaluation nachgelagerten Prozesse eingebunden und/oder haben auch nicht immer Kenntnis von diesen.

Es gibt in Deutschland für öffentliche Institutionen bisher allerdings keine Selbstverpflichtung, mit Evaluationsergebnissen „umzugehen“ und diese aktiv in die Programmsteuerung einzubeziehen. Um sicherstellen zu können, dass Evaluationen auch einen Nutzen bringen, empfehlen Baykal und Kolleg:innen in diesem Zusammenhang, dass

„Ministerien, Förderinstitutionen und durchführende NGOs gleichermaßen [...] moderne Verfahren für die Verwertung von durch Evaluierungen generierten Erkenntnissen und Empfehlungen einführen und selbst Rechenschaft für die Einhaltung ihrer Selbstverpflichtungen ablegen [sollten]“ (vgl. Baykal et al. 2021: 86).

Bei der bundesweiten Befragung haben sich vor allem Evaluationsakteure der Selbstevaluation an Evaluationsergebnissen anderer Projekte und Maßnahmen interessiert gezeigt. Relevant für die Nutzung ist daher auch, ob Evaluationsergebnisse öffentlich zugänglich gemacht werden (können), um diese organisationsübergreifend für die Weiterentwicklung der Praxisarbeit verfügbar zu machen, aber auch, um die Öffentlichkeit über Präventionsbemühungen zu informieren.

Insgesamt zeichnen die Befragungsergebnisse aber ein eher uneinheitliches Bild. So gaben zwar 73% ( $n = 27$ ) der fremdevaluierenden Einrichtungen an, dass die Ergebnisse veröffentlicht werden, im Gegensatz dazu ist dies aber bei nur 49% ( $n = 33$ ) aller selbstevaluierenden Einrichtungen der Fall. Beispielsweise wird dann von einer Veröffentlichung abgesehen, wenn Evaluationen missbräuchlich verwendet werden oder sie Lernprozesse innerhalb einer Organisation verhindern könnten. Auch wenn die Gefahr besteht, dass Personen geschädigt oder gefährdet werden könnten, sollte aus Sicht der Evaluierenden auf eine Veröffentlichung verzichtet werden. In solchen Fällen können der breiten Öffentlichkeit aber Auszüge aus Evaluationen zur Verfügung gestellt werden. Als Alternativen wurden hier z.B. wissenschaftliche Reports, (niedrigschwellige) Blogbeiträge, praxisrelevantes Lehr- und

---

6 Mehrfachantworten waren möglich.

Lernmaterial oder auch Podcasts vorgeschlagen. Dadurch könne der Datenschutz gewahrt werden und ein öffentlicher Austausch stattfinden, auch wenn der Detailgrad dadurch etwas unschärfer sei. Durch die Beauftragungs- und daran anschließende Freigabeprozesse obliegt es aber so gut wie immer den Auftraggebern, ob Evaluationsergebnisse veröffentlicht werden bzw. was mit ihnen geschieht und ob Evaluationen somit ihren Aufklärungs- und Lernfunktionen entsprechen können. Auch fehlt es bislang noch an systematischen Entscheidungskriterien für die Nutzung von Evaluationsergebnissen, sodass die Einrichtungen eher eigene Wege gehen, um ihre Kapazitäten im Hinblick auf die eigene Präventionsarbeit durch Lernen zu verbessern.

Der Austausch zu Evaluationsergebnissen zwischen Trägern in offenen Runden wurde in den Telefoninterviews der bundesweiten Befragung dabei als gewinnbringend betrachtet und könnte zu einer weiteren Expertisenbildung bei den Trägern und damit auch zu einer Stärkung ihrer Evaluationskapazitäten beitragen. Solche Austauschrunden könnten z.B. über unabhängige Einrichtungen ohne Einbezug der Mittelgeber institutionalisiert werden. In den Telefoninterviews konnte insgesamt zwar eine relativ gering ausgeprägte Skepsis gegenüber Evaluationen bzw. Evaluierenden festgestellt werden und oft wurde auch die positive Zusammenarbeit mit den Mittelgebern herausgestellt. Am Ende stand jedoch auch immer die Befürchtung im Raum, bei schlechten Evaluationsergebnissen nicht weiter gefördert zu werden. Dies deutet u.a. auf eine noch verbesserungswürdige Fehlerkultur hin.

Eine positive Fehlerkultur wird in der wissenschaftlichen Auseinandersetzung als eine grundlegende Voraussetzung gesehen, um sinnvolle Präventionsstrategien entwickeln und weiterentwickeln zu können. Auch sollte ein „erfolgreiches Scheitern“ laut den Befragten stärker als Lernfaktor einbezogen werden. So wurde in den Interviews angemerkt, dass es besser wäre, wenn es weniger Evaluationen mit dem Charakter von Öffentlichkeitsarbeit gäbe und nicht der Druck bestünde, sich in gutem Licht präsentieren zu müssen, um zu belegen, dass Gelder vernünftig verausgabt wurden. Vielmehr sollten Evaluationen systematisch zur Neukonzeption von Programmen und Projekten genutzt werden. Nach Ansicht eines Akteurs der Fremdevaluation würden Evaluationen ansonsten auf eine Symbolfunktion reduziert und die Lernfunktion trete in den Hintergrund.

Anhand der Interviewdaten lässt sich an dieser Stelle also zunächst festhalten, dass sich trotz der zunehmenden Nutzung von Fremd- und Selbstevaluationen in zivilgesellschaftlichen Organisationen derzeit in Deutschland noch keine übergreifenden Strukturen und Prozesse etablieren konnten. Dies stellen auch Stockmann et al. in ihrer Untersuchung zur Institutionalisierung von Evaluation in Europa heraus (2020: 192). Evaluationen erfolgen demnach in Deutschland zumeist auf Druck, der von staatlicher Seite auf zivilgesellschaftliche Organisationen ausgeübt wird (Stockmann et al. 2020: 186) – dadurch stehen bei Evaluationen oft Legitimitätszwecke im Vordergrund. Die Aufklärungsfunktion tritt dabei eher in den Hintergrund, was ggf. die Akzeptanzprobleme von Evaluationen verschärfen und die Nachvollziehbarkeit von politischen Entscheidungen erschweren kann (siehe hierzu Stockmann 2010).

Betrachtet man die Förderstrukturen in Deutschland, stellt man zudem fest, dass die Kapazitätensteuerung in der Extremismusprävention und der Demokratieförderung zu einem großen Teil über Evaluationsanteile erfolgt, die in Bundes- und Landesprogrammen verpflichtend angelegt sind. Damit

spielen die Zuwendungsgeber eine zentrale Rolle bei der Schaffung struktureller Voraussetzungen für Evaluation und bei der Weiterentwicklung von Evaluationskapazitäten.

## 4.2 ORGANISATIONSEBENE

Die Mittelhoheit auf Seiten der Programmsteuerung hat somit auch einen großen Einfluss auf die Evaluationskapazitäten auf der Organisationsebene. Diese haben sich zu großen Teilen auch aufgrund des spezifischen Nachfragebedarfs der Förderstrukturen im Bereich der Extremismusprävention und Demokratieförderung herausgebildet und sollen nachfolgend genauer betrachtet werden.

### 4.2.1 ART DER EVALUIERENDEN EINRICHTUNG

#### EINRICHTUNGEN IM BEREICH DER FREMDEVALUATION

Im Bereich der Fremdevaluation wird der überwiegende Teil der Evaluationskapazitäten von außeruniversitären Forschungseinrichtungen mit 32% ( $n = 12$ ), universitären Einrichtungen mit 24% ( $n = 9$ ) und gewinnorientierten Unternehmen mit ebenfalls 24% ( $n = 9$ ) gestellt. Zivilgesellschaftliche und staatliche Träger sowie Stiftungen sind nur zu einem geringen Teil vertreten (Abb. 2).

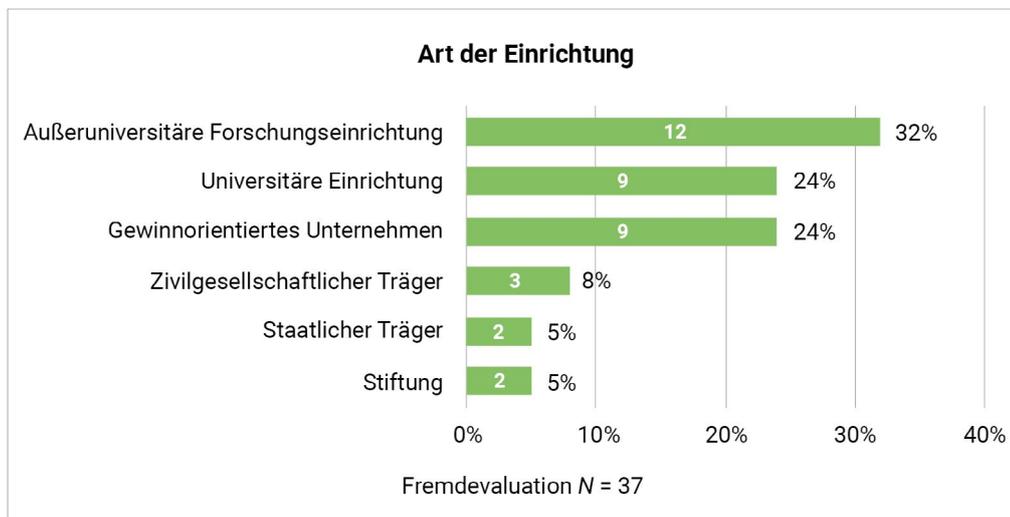


Abb. 2: Häufigkeitsverteilung nach Art der Einrichtung im Bereich der Fremdevaluation.

Evaluation stellt aber bei lediglich 38% ( $n = 14$ ) der extern evaluierenden Einrichtungen den Tätigkeitsschwerpunkt dar. Bei Einrichtungen, bei denen Evaluation nicht den Arbeitsschwerpunkt bildet (62%,  $n = 23$ ), gab es die häufigsten Nennungen bei Forschung mit 35% ( $n = 13$ ), Beratung mit 16% ( $n = 6$ ), Soziale Arbeit mit 3% ( $n = 1$ ) und Sonstiges mit 11% ( $n = 4$ ) als eigentliche/primäre Arbeitsbereiche.

Aufgeschlüsselt nach den unterschiedlichen Einrichtungsarten ergibt sich dabei die in Abbildung 3 dargestellte Verteilung.

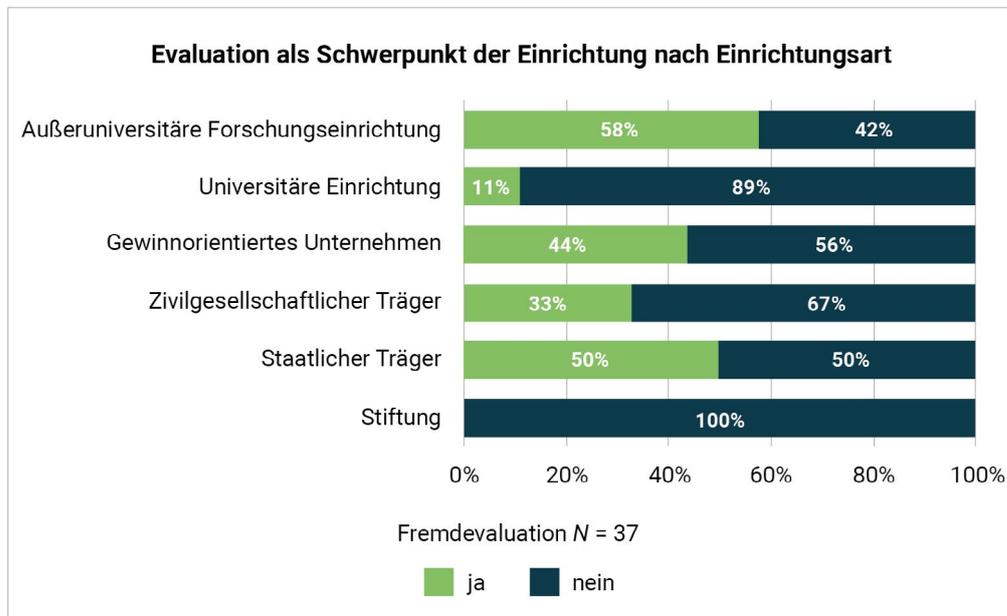


Abb. 3: Evaluation als Schwerpunkt der Einrichtung nach Einrichtungsart im Bereich der Fremdevaluation.

Aus dieser Verteilung lassen sich auch die Expertisenbereiche der unterschiedlichen Einrichtungen ableiten. So kommt universitären und außeruniversitären Forschungseinrichtungen bislang vor allem bei der wissenschaftlichen Begleitung als Sonderform der Evaluation eine wichtige Bedeutung zu. Plausibel ist dies insofern, als wissenschaftliche Begleitungen Wissen über das untersuchte Handlungsfeld erzeugen, Klärungsprozesse anstoßen, Arbeitsprozesse begleiten, zur fachlichen Weiterentwicklung von (Modell-)Projekten beitragen und ggf. auch Bewertungen vornehmen können (vgl. Ehnert et al. 2021: 10). Im Bereich der Evaluationsforschung könnten universitäre Einrichtungen jedoch noch stärker in die Aufgabe der Methodenentwicklung und Weiterentwicklung von Ansätzen eingebunden werden. Ein systematischer Wissenstransfer von der Wissenschaft in die Evaluationspraxis könnte so zu einer organisationsübergreifenden Expertisenbildung und somit zur Stärkung der Evaluationskapazitäten beitragen.

Die Wissenschaft unterstreicht in diesem Zusammenhang immer wieder die Wichtigkeit von unabhängiger Grundlagenforschung. Wie die bundesweite Befragung zeigt, ist diese bisher aber noch nicht ausreichend institutionalisiert und würde auch die aktuellen Evaluationskapazitäten im universitären Bereich überfordern, da weder ausreichend Personal noch finanzielle Mittel zur Verfügung stehen. Daher evaluieren die meisten der interviewten Einrichtungen eher situativ auf Anfrage, d.h., wenn konkrete Bedarfe an sie herangetragen werden. Zumeist gibt es daher für Evaluationen auch

keine festen Mitarbeiter:innen bzw. speziellen Stellen. Zusätzlich bündeln sich Evaluationsexpertisen meist bei einigen wenigen Personen, die beim Verlassen der Universität auch als Kapazität entfallen, was sich anhand gescheiterter Anbahnungsgespräche bei der Befragung auch immer wieder feststellen ließ. Die in universitären Einrichtungen, aber auch in außeruniversitären Forschungseinrichtungen häufig anzutreffende Befristung von Stellen verschärft diese Problematik zusätzlich. Eine auf Langfristigkeit angelegte finanzielle Basisausstattung der Einrichtungen bzw. der Ausbau von Evaluation in Forschung und Lehre könnten hier für eine bessere Planbarkeit und eine Kapazitätserhöhung sorgen, aus der sich auch nachhaltigere Strukturen entwickeln könnten.

Nicht unerwähnt bleiben sollte aber, dass auch nicht forschend tätige Einrichtungen einen Beitrag zur Forschungsarbeit leisten. Einige der gewinnorientierten Unternehmen sind in diesen Bereichen aktiv bzw. entwickeln stetig auch eigene Ansätze weiter. Bei den meisten dieser interviewten Einrichtungen stellt die Forschungstätigkeit aber nicht den Schwerpunkt dar, in diesen Fällen wurden in der bundesweiten Befragung vor allem Evaluations- und Beratungstätigkeiten als Schwerpunkte genannt.

#### EINRICHTUNGEN IM BEREICH DER SELBSTEVALUATION

Im Bereich der Selbstevaluation wurden auf der Ebene von Organisationen in erster Linie zivilgesellschaftliche Träger interviewt (vgl. Abb. 4). Diese stellen eine tragende Säule innerhalb der Extremismusprävention und Demokratieförderung dar (siehe hierzu auch Lützing et al. 2020 und Freiheit et al. 2021b).

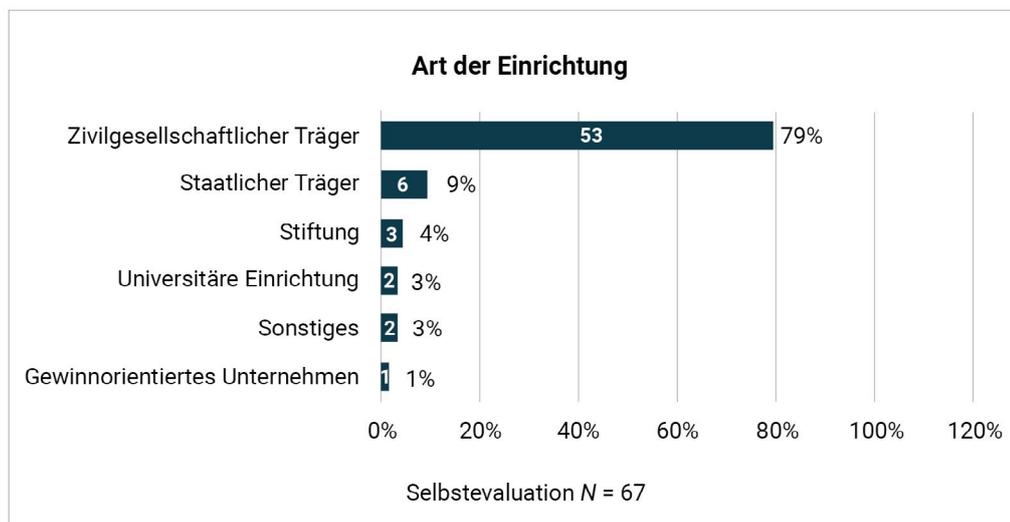


Abb. 4: Häufigkeitsverteilung nach Art der Einrichtung im Bereich der Selbstevaluation.

Aufgrund der eingangs erwähnten doppelten Rollenanforderung von Selbstevaluierenden ist es wichtig, Rollen und Verantwortlichkeiten innerhalb der eigenen Organisation zu klären. Auch müssen in Übereinstimmung mit dem Kriterienkatalog der *DeGEval* die für die Selbstevaluation benötigten per-

sonellen und finanziellen Mittel identifiziert und verfügbar sein (DeGEval 2004). Ansonsten kann sich allein aufgrund der Arbeitsauslastung eine Verschiebung von der Evaluations- hin zur Praxisarbeit vollziehen, was auch in den Interviews der bundesweiten Befragung immer wieder als problematisch herausgestellt wurde.

Deshalb ist speziell für die Selbstevaluation auch immer der jeweilige Anwendungsfall zu betrachten, in dem diese zum Einsatz kommen soll. Vor allem als Mittel zur kontinuierlichen Verbesserung der eigenen Arbeit und als Teil des Qualitätsmanagements können Selbstevaluationen Fremdevaluationen sinnvoll ergänzen (siehe z.B. auch Möller et al. 2020: 403).

#### 4.2.2 EVALUATIONSGEGENSTÄNDE

Neben der Frage, wer evaluiert, ist zur Bestimmung der in Deutschland vorhandenen Evaluationskapazitäten auch die Frage zentral, was evaluiert wird. Nur so kann etwa bestimmt werden, ob die Evaluationsgegenstände von den vorhandenen Einrichtungen auch bestmöglich evaluiert werden können oder ob hier ggf. noch Optimierungsbedarf besteht.

Mit Blick auf die bundesweite Befragung zeigt sich, dass Evaluationskapazitäten vor allem im Bereich der Projekte und Maßnahmen gebunden werden (siehe Abb. 5).

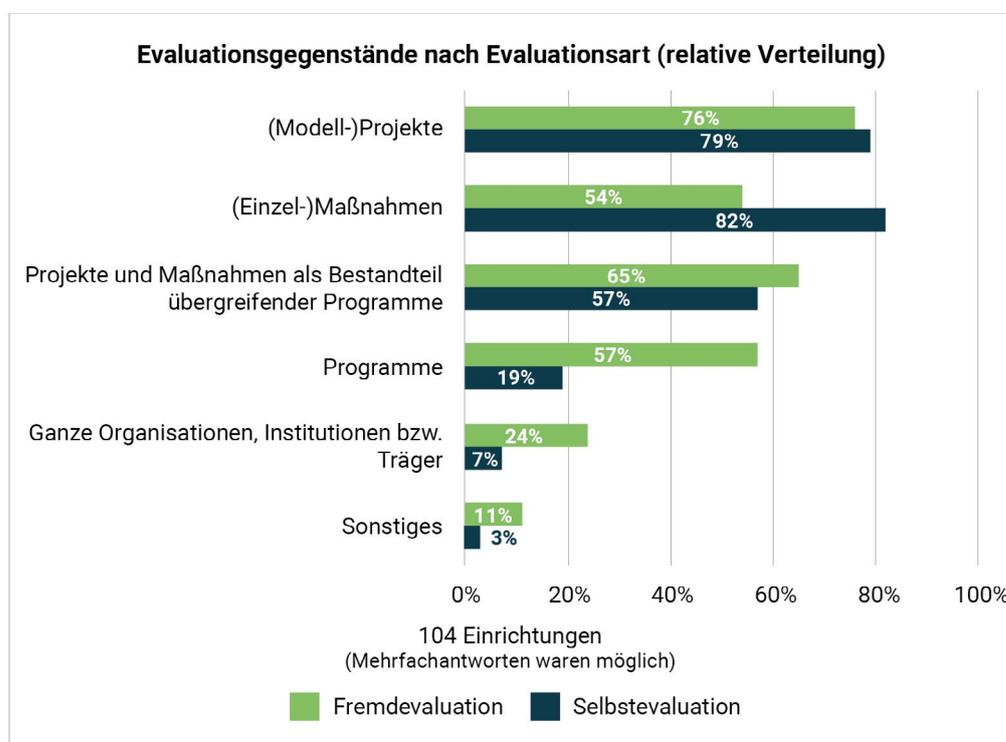


Abb. 5: Relative Verteilung der Evaluationsgegenstände bezogen auf Fremd- und Selbstevaluation.

Wird hierzu die relative Verteilung<sup>7</sup> der Evaluationsgegenstände bezogen auf Fremd- und Selbstevaluationen betrachtet (Mehrfachantworten waren möglich), so zeigt sich, dass (Einzel-)Maßnahmen häufig selbstevaluiert werden. Hier dürften die geringeren Budgetaufwendungen für Selbstevaluationen eine große, wenn nicht die entscheidende Rolle spielen. Projekte und Maßnahmen als Bestandteile von übergreifenden Programmen werden hingegen eher fremdevaluiert. Vor allem bei den großen Bundesprogrammen sind Evaluationen bzw. eine wissenschaftliche Begleitung in den Programmen bereits mit angelegt, deren Durchführung oft auch verpflichtend für die Zuwendungsempfänger ist. Im Bundesprogramm *Demokratie leben!* ist beispielsweise sowohl eine Evaluation auf Gesamtprogramm- als auch Projektebene vorgesehen (BMFSFJ 2020; BMFSFJ 2021).

Überraschend ist in diesem Zusammenhang der relativ hohe Anteil selbstevaluierender Einrichtungen, die in den Interviews angaben, auch Programme zu evaluieren (19%;  $n = 13$ ). Diese werden allerdings meist vom Träger selbst umgesetzt oder es handelt sich um Programme mit einem geringeren Finanzvolumen, die meist von Länderebene finanziert und über einen Träger koordiniert werden. Dies sind beispielsweise Förderprogramme im Bereich der mobilen Jugendarbeit, bei denen kleinere Summen an Initiativen vor Ort vergeben werden. Die Selbstevaluation erfolgt in solchen Fällen dann mit den jeweils eigenen Kapazitäten innerhalb des koordinierenden Trägers, nur eben auf Programmebene. Ob dabei der Rahmen einer Selbstevaluation angemessen ist, kann hier nicht abschließend beurteilt werden. Anhand der Daten der bundesweiten Befragung lässt sich allerdings feststellen, dass Programmevaluationen üblicherweise von fremdevaluierenden Einrichtungen durchgeführt werden. Einerseits sind hier eine spezifische Evaluationsexpertise und -methoden erforderlich, andererseits werden solche Programme aber auch meist von Einrichtungen des Bundes oder Landes initiiert, denen keine eigenen internen Evaluationskapazitäten zur Verfügung stehen und die Evaluationen daher an externe Einrichtungen vergeben.

Auch im Bereich der Organisationsevaluation finden eher klassische externe Beratungsdienstleistungen im Rahmen einer Organisationsberatung bzw. Organisationsentwicklung statt. Selbstevaluationen der eigenen Organisation im Sinne eines laufenden Qualitätsmanagements wurden in der Befragung insgesamt seltener genannt. So sind die Kapazitäten in diesem Bereich bislang vor allem bei fremdevaluierenden Einrichtungen verortet (vgl. Abb. 5).

#### 4.2.3 EVALUATIONSFORMEN

Evaluationsaufträge können zumeist erst dann vergeben werden, wenn Projekte bewilligt bzw. mit Finanzmitteln ausgestattet wurden. So können Evaluationen aus strukturellen Gründen oft erst während der Projekt- oder Maßnahmendurchführung ansetzen.

---

7 Die relative Verteilung stellt die Prozentangaben bezogen auf die jeweilige Stichprobengröße dar. Absolute Vergleiche zwischen Fremd- und Selbstevaluation sind aufgrund der unterschiedlichen Stichprobenziehung und -beschaffenheit nicht zulässig. Beide Perspektiven sollten demnach auch immer getrennt voneinander betrachtet werden, da sie ihren jeweils eigenen Handlungslogiken folgen und spezifischen Herausforderungen begegnen.

Dies wurde bei der Befragung vor allem von den extern evaluierenden Einrichtungen kritisiert. Ein früher Beginn von Evaluationen und phasenübergreifende, d.h. begleitende, Evaluationen werden hingegen als besonders hilfreich hervorgehoben, um zur Verbesserung von Projekten oder Programmen beizutragen. In der Realität lässt sich dies aber oft nicht einlösen (siehe Abb. 6). Eine Evaluatorin bemängelte in diesem Zusammenhang etwa, dass durch lange Ausschreibungsprozesse für Evaluationsaufträge wertvolle Zeit verloren geht. In Kombination mit den begrenzten Laufzeiten von Modellprojekten setze Evaluation dann zu spät an und es bliebe zu wenig Zeit für die eigentliche Evaluation.

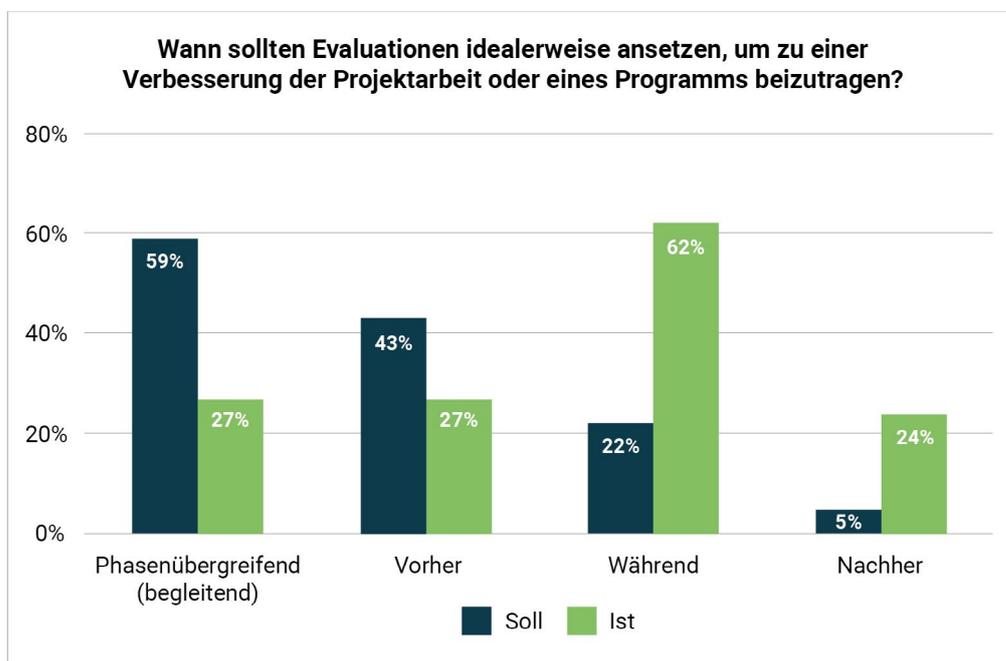


Abb. 6: Soll-Ist-Vergleich des Evaluationsbeginns im Bereich der Fremdevaluation (der besseren Lesbarkeit wegen werden die Antworten „keine Angabe“ und „weiß nicht“ hier nicht dargestellt).

Im Bereich der Selbstevaluation sind die Handlungsspielräume flexibler, da die Bereitstellung von Ressourcen keine gesonderten Beauftragungsprozesse für Externe erfordert. Auch hier können Evaluationen jedoch nicht immer rechtzeitig ansetzen. So gaben die selbstevaluierenden Einrichtungen an, dass Evaluationen, die vor Beginn der Projekt- oder Maßnahmendurchführung ansetzen, einen erhöhten Personalaufwand darstellen würden, der sich nicht immer abdecken ließe.

Diese sogenannten *Ex-ante*-Evaluationen stellen im Anschluss an Silvestrini aber eine wesentliche Voraussetzung dar, um – die von Mittelgeberseite oft (zusätzlich) gewünschten – Wirkungsanalysen durchführen zu können, Programme besser steuern und Risiken für nicht-intendierte Wirkungen verringern zu können (Silvestrini 2016: 269–281).

An dieser Stelle scheinen sich auch die Perspektiven der Evaluationsakteure von denen der Präventionspraxis zu unterscheiden. So gaben im Rahmen der Bedarfserhebung von Koynova und Kolleg:innen nur 22% der befragten Praxisakteure *Ex-ante*-Evaluationen als gewünschte Evaluationsform an (siehe Infobox).

#### EVALUATIONSFORMEN: PERSPEKTIVEN AUS DER PRÄVENTIONSPRAXIS

- ▶ 61 % aller Befragten äußern einen Bedarf an begleitenden Evaluationen
- ▶ 52 % an Evaluationen, die rückblickend durchgeführt werden
- ▶ 22 % an Evaluationen, die vor oder zu Beginn der Durchführung von Maßnahmen stattfinden

Koynova et al. 2022: 24–25.

Auch zeigten sich die befragten Präventionsakteure in vertiefenden Interviews und Workshops vielmehr „überrascht, dass die Weichen für die Monitoring- und Evaluationsprozesse bereits zu Beginn eines Projektes gestellt werden. *Ex-ante*-Evaluationen waren entsprechend auch eher unbekannt“ (Koynova et al. 2022: 24). Es wurde aber gleichzeitig ein Mehrbedarf an Kontext- und Risikoanalysen sowie Relevanzuntersuchungen geäußert, um überprüfen zu können, ob Bedarfe und Angebote zueinander passen (Koynova et al. 2022: 24). Diese Formate erfordern jedoch, dass Evaluationen früh ansetzen. Es lässt sich also dennoch ableiten, dass auch aus Sicht der Praxisakteure ein früher Einbezug bzw. Beginn von Evaluationen wichtig wäre, um eine Verbesserung präventiver und demokratiefördernder Bemühungen zu erreichen.

Begleitende Evaluationen wurden hingegen sowohl von evaluierender als auch evaluierter Seite begrüßt (siehe Infobox und Abb. 6). Da diese aber ebenfalls zu einem erhöhten Zeit- und Personalbedarf führen, wäre auch hier, wie die befragten Einrichtungen im Bereich der Fremdevaluation betonen, eine Verbesserung der entsprechenden strukturellen Rahmenbedingungen zur Stärkung der Evaluationskapazitäten notwendig und wünschenswert.

#### 4.2.4 LOKALE VERORTUNG

Weiterhin hat sich in den Interviews der bundesweiten Befragung gezeigt, dass die örtliche Nähe der Evaluierenden zu den Evaluierten eine nicht zu vernachlässigende Rolle spielt. Sie hilft beispielsweise, eine ggf. vorhandene Skepsis der Evaluierten durch einen intensiveren persönlichen Kontakt abzubauen und/oder Termine vor Ort zeitlich flexibler und kostengünstiger realisieren zu können.

In der Literatur findet die lokale Verortung von Evaluationskapazitäten zumeist eher wenig Beachtung. In diesem Zusammenhang sind besonders die extern evaluierenden Einrichtungen als die evaluationsdurchführenden Organisationen relevant. Wie auf der Kapazitätenübersicht der online zu-



In Abbildung 8 sind die Präventionsangebote im Bereich des religiös begründeten Extremismus bzw. islamistischen Extremismus grau eingefärbt, die externen Evaluationsangebote in diesem Bereich rot, Kreise fassen mehrere Angebote bzw. Einrichtungen zu Clustern zusammen. In den Ballungszentren in Nordrhein-Westfalen und in den bereits genannten Bundesländern ohne Evaluationsangebote sind insgesamt nur wenig Evaluationskapazitäten in diesem Phänomenbereich vorhanden. So gibt es in Hamburg beispielsweise aktuell keine Evaluationseinrichtungen in diesem Arbeitsbereich, dem stehen aber 42 Präventionsangebote im Bereich des religiös begründeten Extremismus bzw. islamistischen Extremismus gegenüber (siehe blaue Umrandung in Abb. 8).

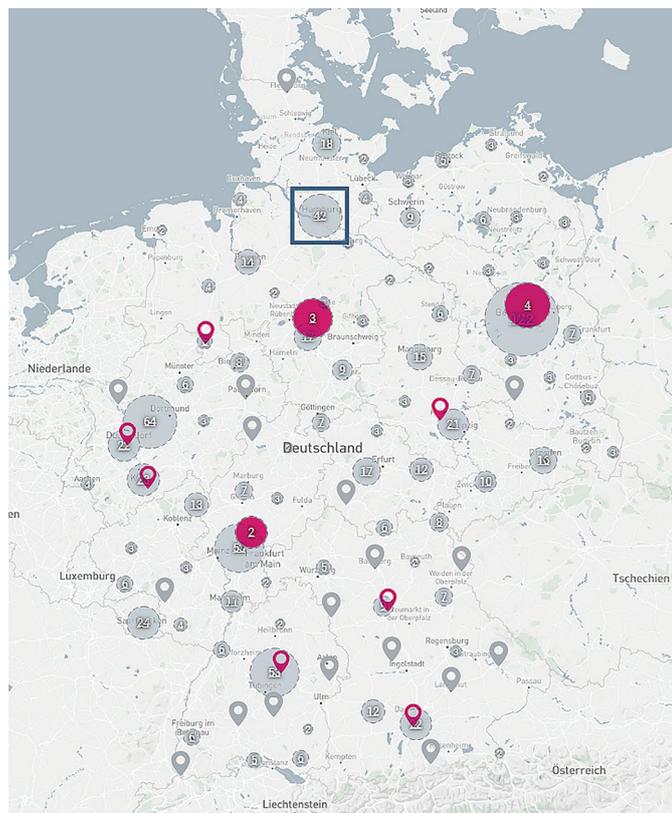


Abb. 8: Externe Evaluationskapazitäten und Präventionsangebote im Bereich des religiös begründeten Extremismus bzw. islamistischen Extremismus (Überlagerung der PrEval- und MAPEX-Daten).

### 4.3 AKTEURSEBENE

Während sich den Evaluationskapazitäten bislang auf der strukturellen und Organisationsebene angenähert wurde, soll der nächste Systematisierungsschritt entlang der Akteursebene erfolgen (vgl. Tab. 1).

### 4.3.1 QUALIFIKATIONEN

Die Diversität der deutschen Präventionslandschaft stellt dabei hohe Anforderungen an Evaluat:innen, da Evaluationen immer auch individuell an den jeweiligen Evaluationsgegenstand angepasst werden müssen, um aussagekräftige Ergebnisse erzielen zu können (siehe z.B. Nehlsen et al. 2020: 2).

#### QUALIFIKATIONEN IM BEREICH DER FREMDEVALUATION

Gefragt nach besonderen Kenntnissen bzw. Qualifikationskriterien, die Mitarbeiter:innen für die Evaluationstätigkeit mitbringen müssen, benannten 78% ( $n = 29$ ) der extern evaluierenden Einrichtungen Methodenkenntnisse als wichtigste Einstellungsvoraussetzung. Auch eine gewisse Praxiserfahrung sei einstellungsrelevant und wurde von 30% ( $n = 11$ ) aller Befragten angegeben. Erste Erfahrungen in der Projekt- bzw. Evaluationsarbeit spielen hingegen mit 8% ( $n = 3$ ) eine eher untergeordnete Rolle. Sozial- und Selbstkompetenzen<sup>9</sup> wurden von 16% ( $n = 6$ ) der befragten Einrichtungen als wichtig herausgestellt (siehe Abb. 9). Diese Qualifikationsbereiche (Methodenkenntnisse, Praxiserfahrung, Sozial- bzw. Selbstkompetenzen) finden sich auch in den von der *DeGEval* empfohlenen Kompetenzfeldern und Dimensionen für die Aus- und Weiterbildung im Bereich der Evaluation wieder (DeGEval 2008). Sie sind dementsprechend für die Evaluationsdurchführung besonders relevant.

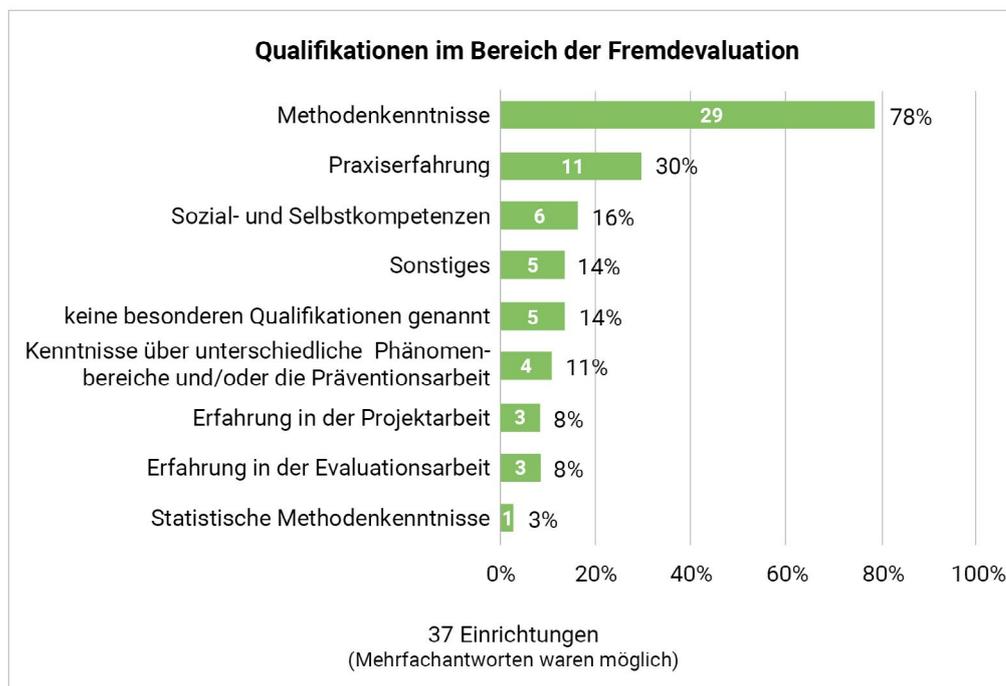


Abb. 9: Qualifikationen im Bereich der Fremdevaluation.

<sup>9</sup> Diese Kompetenzen beziehen sich auf die Empfehlungen der *DeGEval* für die Aus- und Weiterbildung in der Evaluation. Dimensionen der Sozial- bzw. Selbstkompetenzen sind dabei soziale und kommunikative Kompetenz, Kooperationskompetenz, Selbstmanagementkompetenz, Lern- und Problemlösekompetenz (DeGEval 2008).

Sachkenntnisse werden zwar von der Fachpraxis, also den Evaluierten, als besonders wichtig herausgestellt (siehe Infobox), im Rahmen der Kapazitätenerhebung gaben allerdings nur 11% ( $n = 4$ ) der befragten fremdevaluierenden Einrichtungen an, dass phänomenspezifische Kenntnisse oder Kenntnisse über die Präventionsarbeit erforderliche Qualifikationen bei der Einstellung neuer Mitarbeiter:innen wären.

### QUALIFIKATIONEN: PERSPEKTIVEN AUS DER PRÄVENTIONSPRAXIS

In den Bedarfserhebungen des *PrEval*-Forschungsverbunds im Bereich der Präventionspraxis wurde vonseiten der Evaluierten die Sachkenntnis von Evaluierenden als wichtiges Kriterium gelingender Evaluation herausgehoben, um anwendungsorientierte Empfehlungen bzw. nachgelagerte Lernprozesse bei den Evaluierten anzustoßen:

„Einige Befragte glaubten, das sei nur dann möglich, wenn Evaluierende selbst über Sachkenntnis verfügen und auch die Praxisperspektive auf den Arbeitsbereich kennen“ (Koynova et al. 2022: 28).

Vor diesem Hintergrund stellt sich die Frage, inwieweit ein Umdenken bzw. eine themen- und phänomenspezifische Ausbildung notwendig ist, um den oben genannten Bedarfen der Evaluierten besser begegnen zu können. Aus betriebswirtschaftlichen Gründen ist es zwar nachvollziehbar, dass sich gewinnorientierte Unternehmen so breit wie möglich aufstellen, um nicht nur „spartenspezifische“ Evaluationen durchführen zu können. Auch gab ein Interviewpartner der Fremdevaluation an, dass bei einer entsprechenden engen Zusammenarbeit mit den zu evaluierenden Projekten fehlende Feld- bzw. Sachkenntnisse kompensiert werden könnten. Es bleibt aber noch zu klären, ob ein weiterer Kapazitätenausbau für das hochspezifische Feld der Extremismusprävention notwendig ist.

Im Vergleich zu anderen Ländern, etwa der Schweiz, gibt es in Deutschland generell weniger Ausbildungsangebote, die speziell auf Evaluation ausgerichtet sind. In der universitären Ausbildung konnten in einer bundesweiten Studie von 2017 in nur zwei Fällen spezielle Evaluationsmodule in allen untersuchten Studiengängen gefunden werden (Stockmann et al 2020: 187–188). Evaluation ist in den meisten Fällen also eher eine Tätigkeit, die in Deutschland *on the job* gelernt wird (Stockmann et al. 2020: 188), d.h., die Kenntnisse bzw. Kapazitäten werden bislang zumeist in der Berufspraxis ausgebildet.

### QUALIFIKATIONEN IM BEREICH DER SELBSTEVALUATION

Im Bereich der Selbstevaluation hat die Weiterbildung und laufende Qualifizierung der für die Evaluationstätigkeiten eingesetzten Mitarbeiter:innen einen besonderen Stellenwert. Da Projektmitarbei-

tende die Evaluationen selbst durchführen, wird auf die jeweils intern vorhandenen oder ggf. bereits herausgebildeten Kapazitäten zurückgegriffen.

Bei der Frage, welche besonderen Kenntnisse bzw. Qualifikationskriterien für die Evaluationstätigkeit mitgebracht werden müssen, nannten die meisten der Befragten dementsprechend auch keine speziellen Qualifikationserfordernisse (36%;  $n = 24$ ).

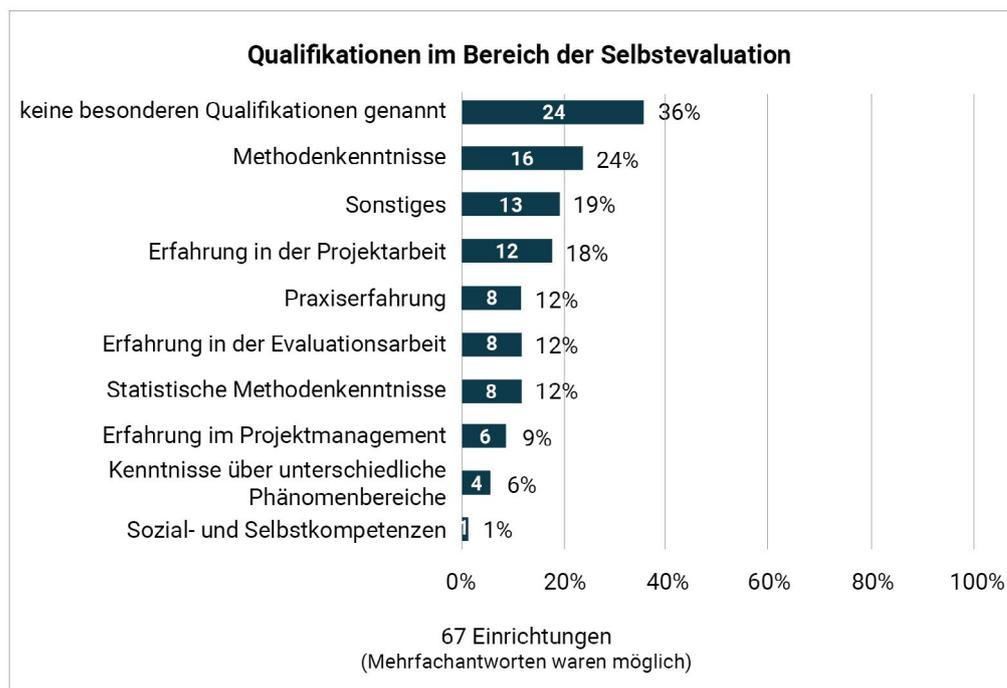


Abb. 10: Qualifikationen im Bereich der Selbstevaluation.

Waren Qualifikationen hingegen erforderlich, wurden Methodenkenntnisse mit 24% ( $n = 16$ ), Erfahrung in der Projektarbeit mit 18% ( $n = 12$ ), Praxiserfahrung, Erfahrung in der Evaluationsarbeit und statistische Methodenkenntnisse mit jeweils 12% ( $n = 8$ ) angegeben. Bestimmte Voraussetzungen für die Evaluationstätigkeit im Sinne des *DeGEval*-Kriterienkatalogs für die Selbstevaluation werden also bereits erfüllt und deuten auf einen gewissen Professionalisierungsgrad in diesem Bereich hin.

Zum Teil begleiten und beraten aber auch fremdevaluierende Einrichtungen die befragten selbst-evaluierenden Projektträger bei der Umsetzung und Durchführung von Evaluationen, so beispielsweise bei der Entwicklung von Fragebögen. Die dadurch aufgebauten Kapazitäten bzw. Evaluationsexpertisen könnten somit für zukünftige Evaluationen besser nutzbar gemacht werden (vgl. hierzu auch Brinkmann et al. 2019). Vor allem bei auf Dauer angelegten Kooperationen könnten sich Fremd- und Selbstevaluationen gegenseitig stärken. Dadurch ließen sich Feldkenntnisse auf Seiten fremdevaluierender und Methodenkenntnisse auf Seiten selbstevaluierender Einrichtungen ausbilden und verfestigen.

Des Weiteren könnte das in Evaluationen generierte Wissen in der Zusammenarbeit mit universitären Einrichtungen auch stärker in die universitäre Ausbildung zurückfließen und so zusätzliche Ausbildungsformen bzw. gezielt auf Evaluation (in der Extremismusprävention) ausgerichtete Studiengänge herausbilden.

#### 4.3.2 EVALUATIONSANSÄTZE UND -MODELLE

Um zu aussagekräftigen Evaluationsergebnissen gelangen zu können, ist unter anderem die Wahl geeigneter Evaluationsansätze und -modelle ausschlaggebend und auf Seiten der Evaluationsakteure ein breites Methodenwissen vorteilhaft.

##### EVALUATIONSANSÄTZE UND -MODELLE IM BEREICH DER FREMDEVALUATION

Zunächst lässt sich anhand der bundesweiten Befragung feststellen, dass Fremdevaluationsakteure mit ihrem Methodenwissen und den für die Evaluation genutzten Ansätzen und Modellen breit aufgestellt sind (siehe Abb. 11).

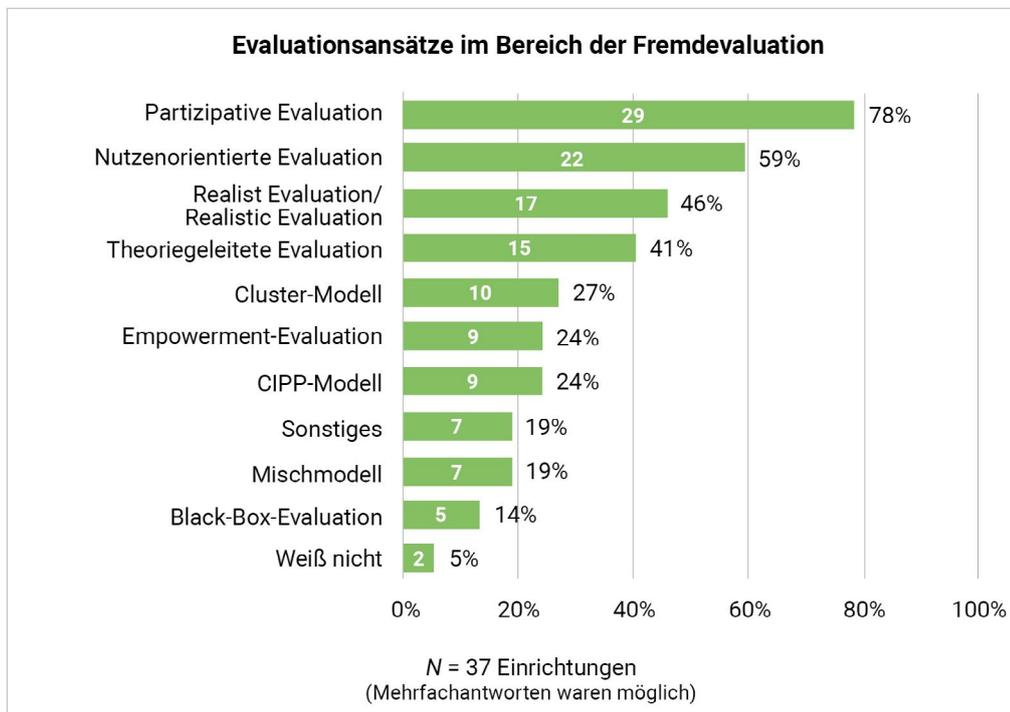


Abb. 11: Evaluationsansätze im Bereich der Fremdevaluation.

Im Bereich der Extremismusprävention und Demokratieförderung werden in der Evaluationsforschung dabei besonders diejenigen Ansätze als vielversprechend diskutiert, bei denen relevante Kontextfaktoren einbezogen werden (Nehlsen et al. 2020: 11–14). Aus Sicht von Präventionsakteuren ist hier vor allem ein dialogisches und partizipatives Vorgehen begrüßenswert (siehe Infobox).

#### EVALUATIONSANSÄTZE UND -MODELLE: PERSPEKTIVEN AUS DER PRÄVENTIONSPRAXIS

„Bei externen Evaluationen sollten die Evaluierenden ein dialogisches und partizipatives Vorgehen anstreben. Externe Evaluationen sollten die Erkenntnisinteressen der Praxis mitberücksichtigen, Bewertungskriterien im Dialog mit der Praxis entwickeln, die Anforderungen des Projektalltags respektieren und besondere Datennutzungsbedingungen berücksichtigen.“

Koynova et al. 2022: 33.

Die Bedeutung des dialogischen Vorgehens und der offenen Gestaltung des Evaluationsprozesses zeigt sich auch in den Daten der bundesweiten Befragung. So gaben 78% ( $n = 29$ ) der fremdevaluierenden Einrichtungen an, partizipative Ansätze bei ihrer Evaluationsarbeit einzusetzen, was sowohl Projektteams als auch am Evaluationsgegenstand beteiligte Stakeholder einschließt. Auch bei der Auswahl und Operationalisierung geeigneter Erfolgskriterien werden alle Projektbeteiligten einbezogen. Nach Aussagen der befragten fremdevaluierenden Einrichtungen sei die Einbindung nicht nur insgesamt lohnenswert, sondern durch die gemeinsame Festlegung konkreter Ziele (vor der Durchführung der eigentlichen Evaluation bzw. Projektarbeit) ließen sich später auch nutzbringende, *Outcome*-basierte Ergebnisse erzeugen.

Das in der Praxis vorhandene Handlungswissen sollte deshalb in der Konzeptionsphase von Evaluationen ausreichend Berücksichtigung finden, um dieses im Rahmen der Evaluation überprüfen und später in die Weiterentwicklung der Präventionsansätze einbeziehen zu können. Auf der Ebene der Strukturen und Organisationen bedarf es hierzu jedoch wiederum ausreichender finanzieller und zeitlicher Ressourcen.

Auch bei nutzenorientierten Evaluationen (59%;  $n = 22$ ) ist der dialogische Einbezug von Stakeholdern wichtig, um die spätere Akzeptanz der Evaluationsergebnisse und den Nutzungswillen bei den involvierten Akteuren zu steigern (siehe hierzu auch Patton 2008) und Lernprozesse anregen zu können.

Zu beachten ist allerdings, dass in der Befragung vor allem diejenigen Ansätze und Modelle von den Interviewten genannt wurden, die durch den Fragebogen vorgegeben waren. Aktuell in der Forschung diskutierte und ebenfalls als vielversprechend geltende Modelle, etwa das *Logische Modell* (siehe hierzu z.B. Beywl et al. 2007), fanden sich nur vereinzelt in den offenen Antworten der befragten Evaluationsakteure wieder. Dies könnte auf Verbesserungspotenziale beim Wissensaustausch

zwischen Forschung und Evaluation hindeuten. Generell wurde in der bundesweiten Befragung der teils langsame Transfer wissenschaftlicher Ergebnisse in die Praxis bemängelt.

Durch eine bessere Verzahnung zwischen Wissenschaft und Praxis könnte das Handlungswissen aus der Praxis schneller in die Forschung zurückfließen. Dadurch könnten Evaluationsansätze mit Handlungswissen aus der Praxis angereichert und überarbeitet werden und so zu einer besseren Passung gelangen. Für die Evaluationskapazitäten würde das eine schnellere und wissenschaftlich fundierte Expertenbildung bedeuten, die zeitlich näher abgestimmt auf Veränderungen von Radikalisierungsphänomenen reagieren könnte. Diese enge Verzahnung benötigt aber sowohl Zeit- als auch Personalressourcen und erfordert damit auch eine Finanzierung, welche oft nicht ausreichend vorhanden und für diese Zwecke nicht vorgesehen ist.

#### **EVALUATIONSANSÄTZE UND -MODELLE IM BEREICH DER SELBSTEVALUATION**

Eine qualitative Einschätzung der in der Selbstevaluation vorhandenen Kenntnisse von Evaluationsansätzen und -modellen lässt sich anhand der Daten der bundesweiten Befragung nur indirekt treffen. So ließ sich in den Telefoninterviews feststellen, dass die gängigsten Ansätze und Modelle (diese wurden den Interviewten vorgelesen) den Befragten zum größten Teil unbekannt waren und eher eigene Ansätze genannt wurden. Oft entschieden sich die interviewten Selbstevaluationsakteure erst für eine der Auswahlmöglichkeiten, nachdem ihnen die verschiedenen Ansätze und Modelle auf Nachfrage näher erläutert wurden. Die ausgewählte(n) Antwort(en) wies(en) dann zumeist die größte Übereinstimmung mit dem eigenen Vorgehen auf. Im Rahmen der Befragung traf dies in erster Linie auf partizipative, nutzenorientierte und Empowerment-Ansätze zu. Ein Interviewpartner gab zudem an, dass die Ansätze eher „intuitiv“ angewendet würden, „ohne das Konzept genau zu verfolgen“. Ein ähnliches Antwortverhalten zeigte sich auch in weiteren Interviews. Es kann also davon ausgegangen werden, dass hier noch Expertenlücken bestehen.

Als gewinnbringend könnten sich auch hier unterstützte Selbstevaluationen durch externe Fachkräfte erweisen. Der gegenseitige Wissensaustausch und -transfer könnte Experten bei den Selbstevaluationskapazitäten stärken (vgl. hierzu DeGEval-Standard N3, DeGEval 2004: 10) und damit auch insgesamt zu einer verbesserten Lernkultur beitragen.

#### **4.4 DIE PREVAL-PLATTFORM – EIN DATENVISUALISIERUNGSTOOL FÜR DIE EVALUATION**

Die hier präsentierten Daten stellen einen ausgewählten Ausschnitt aus den Ergebnissen der bundesweiten Befragung dar. Wichtig für ein gezieltes Monitoring und eine detaillierte Analyse des Evaluationsgeschehens in Deutschland ist eine umfassende und systematische Übersicht über die Evaluationslandschaft. Eine solche Übersicht wurde sowohl im Bereich der Präventions- als auch in der Evaluationsforschung immer wieder gefordert, um einerseits die Vernetzung der einzelnen Akteure zu fördern und andererseits, um eine Bestandsaufnahme der vorhandenen Experten und Kapazitäten abzubilden und diese analysierbar zu machen. Diese Lücke sollte die *PrEval*-Plattform für die

Evaluation im Bereich der Extremismusprävention und Demokratieförderung schließen. Die Plattform setzt sich dabei aus zwei wesentlichen Ansichten zusammen, die unterschiedliche Funktionsebenen abdecken: der Kapazitätenübersicht und der Analyseansicht.

### KAPAZITÄTENÜBERSICHT

In der Kapazitätenübersicht finden sich sämtliche interviewten Akteure im Bereich der Selbst- und Fremdevaluation, die ihr Einverständnis für die Veröffentlichung auf der Landkarte gegeben haben. Zum Zeitpunkt der vorliegenden Auswertung (Stand: 25.10.2021) waren dies:

- 34 Einrichtungen im Bereich der Fremdevaluation
- 63 Einrichtungen im Bereich der Selbstevaluation

Dabei können die verschiedenen Einrichtungen durchsucht und anhand der Evaluationsart, der evaluierten Arbeitsbereiche und Präventionsebenen sowie der Standorte gefiltert werden. Auch können die Einrichtungen entweder direkt auf der Karte oder als Liste angezeigt werden.

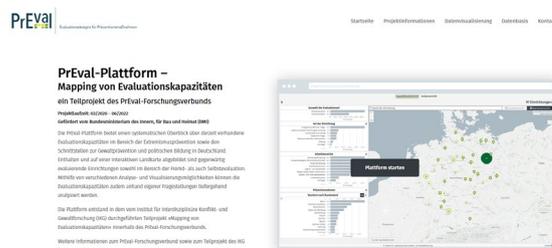
Mit einem Klick auf die Marker in der Kartenansicht oder den Namen der Einrichtung in der Listenansicht öffnet sich zudem ein Fenster mit einer Kurzbeschreibung, der Adresse sowie der Webseite der jeweiligen Einrichtung (Abb. 12).

## DIE PREVAL-PLATTFORM

Das digitale Analyse und Datenvisualisierungstool, das im Rahmen des *PrEval*-Verbundprojekts umgesetzt wurde, baut auf der *MAPEX*-Plattform auf. Diese bildet einen umfassenden Datensatz aus der Präventionslandschaft im Bereich der islamistischen Radikalisierung im Internet öffentlich zugänglich ab.

Die Plattform wird nun um den *PrEval*-Datensatz erweitert und zeigt die Daten sowohl aus Perspektive der Fremd- als auch der Selbstevaluation.

Erreichbar ist sie unter [preval-plattform.de](https://preval-plattform.de).



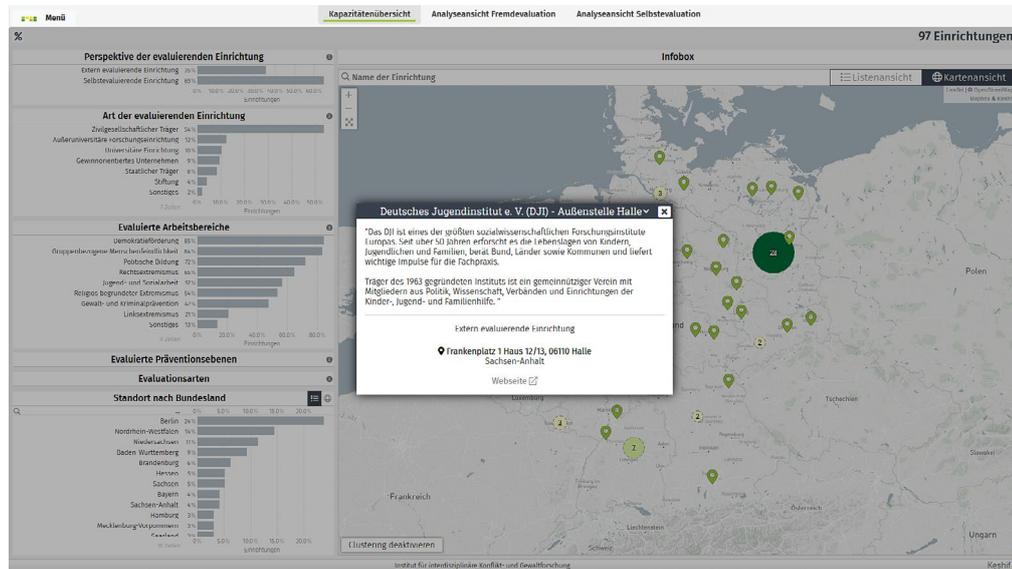


Abb. 12: Die Kartenansicht in der Kapazitätenübersicht.

Im Vordergrund dieser Ansicht stehen folglich Vernetzungsaspekte, auch soll eine schnelle, aber umfassende Übersicht über die Evaluationskapazitäten ermöglicht werden. Anhand der dargestellten Cluster lässt sich bundesweit zudem die Verteilung bzw. Ballung der Evaluationskapazitäten nachvollziehen.

### ANALYSEANSICHTEN

Im Unterschied zur Kapazitätenübersicht unterteilt sich die Analyseansicht in die Bereiche „Fremdevaluation“ (37 Einrichtungen) und „Selbstevaluation“ (67 Einrichtungen). Diesen beiden Analyseansichten (s. Abb. 13) liegt jeweils ein anonymisierter Datensatz zugrunde. Die hier gemachten Angaben sind also nicht mehr auf die einzelnen Einrichtungen zurückführbar.

Auch bietet die Analyseansicht im Unterschied zur Kapazitätenübersicht eine Vielzahl weiterer Analyse- und Visualisierungsmöglichkeiten. So lässt sich anhand der Kategorien beispielsweise nachvollziehen, welche Ansätze und Methoden bei der Evaluation angewendet werden und in welchem Verhältnis diese zueinander stehen. Ebenfalls lassen sich die verschiedenen Evaluationsgegenstände (Projekte, Maßnahmen oder Programme) weitergehend analysieren oder auch darstellen, wie Evaluationsergebnisse später verwendet werden.

Dadurch lassen sich leicht Häufigkeiten ablesen und Beziehungen zwischen einzelnen Erhebungskategorien bilden, die auf Trends hinweisen können, aber auch Schwerpunkte und Lücken erkennen lassen. So kann mithilfe der Analyseansichten beispielsweise aufgezeigt werden, dass Evaluierenden im Bereich der Selbstevaluation insgesamt mit weniger Skepsis begegnet wird als im Bereich der Fremdevaluation (siehe Abb. 14). Das ist insofern auch naheliegend, da Selbsteva-

lationen entweder von den eigenen Teammitgliedern oder einer Person innerhalb der eigenen Organisation durchgeführt werden, während Fremd- bzw. externe Evaluationen durch Personen außerhalb der eigenen Organisation erfolgen, zu denen ggf. das Vertrauen fehlt und erst noch aufgebaut werden muss.

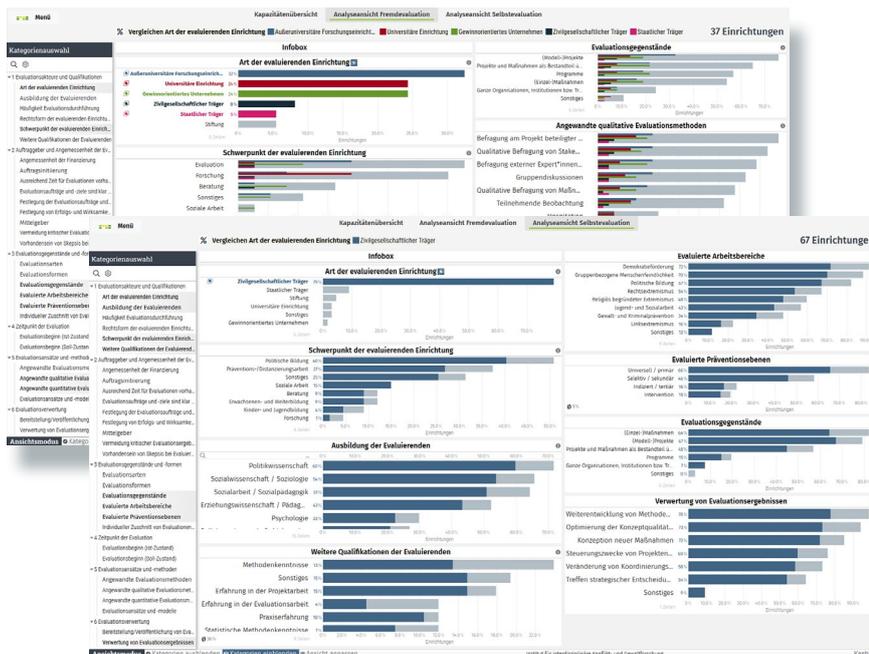


Abb. 13: Die beiden Analyseansichten im Bereich der Fremd- und Selbstevaluation.

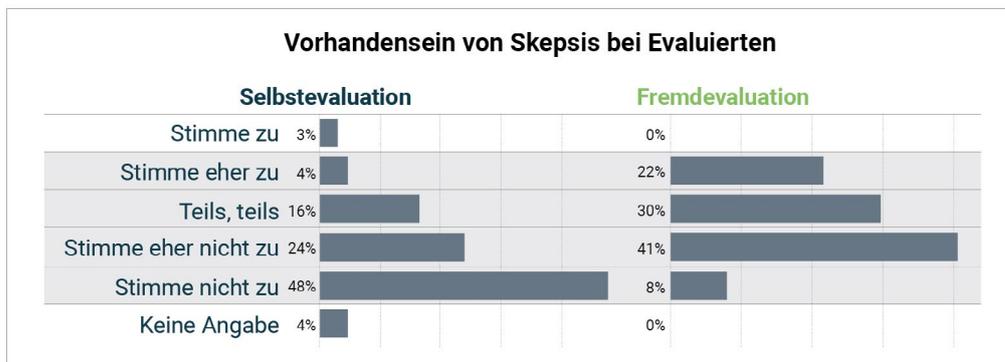


Abb. 14: Skepsis im Bereich der Fremd- und Selbstevaluation bezogen auf die Anzahl der jeweils interviewten Einrichtungen (die Abbildung kombiniert beide Analyseansichten).

Die *PrEval*-Plattform ist zudem offen gestaltet und bietet flexible Schnittstellen, sodass jederzeit weitere Daten integriert werden können. Vor allem in einer systematischen Längsschnittstudie könnten so Aussagen über Veränderungen und Entwicklungen getroffen werden, wodurch überprüfbar würde, ob vormals identifizierte Lücken geschlossen wurden oder neue Herausforderungen hinzugekommen sind. Auch ließen sich Perspektiven zusammenführen und im besten Fall nachhaltige, strukturbildende Maßnahmen ableiten. Diese könnten zu einer Verbesserung der Evaluation und Praxisarbeit im Bereich der Extremismusprävention und Demokratieförderung beitragen.

## 5. INTERNATIONALE VERORTUNG

Wie die vorausgegangenen Analysen gezeigt haben, kommt der strukturellen Ebene insgesamt eine wichtige Aufgabe bei der Ausbildung, Weiterentwicklung und Stärkung von Evaluationskapazitäten zu. Durch die in Deutschland zum überwiegenden Teil von staatlicher Seite bereitgestellten Mittel für Evaluationen übernehmen mittelgebende Einrichtungen neben der Finanzierung bzw. mit der Finanzierung auch eine steuernde Funktion (Stockmann 2008: 5). Folglich betrifft die Entwicklung von Evaluationskapazitäten auch jene Akteure, die Evaluationen verwalten und in Auftrag geben und dementsprechend ein Interesse an Evaluationsergebnissen auf der politischen Ebene und der Programmebene haben.

Im Bereich der Extremismusprävention ist die Evaluation in Deutschland ein noch relativ junges Themenfeld. Bei der Bestimmung und Weiterentwicklung von Evaluationskapazitäten lohnt sich von daher nicht nur ein Blick in andere Handlungsbereiche, wie z.B. der Entwicklungszusammenarbeit, sondern auch in andere Länder und auf die dort bestehenden Evaluationsstrukturen und -praktiken. Die von der Prävention adressierten und von der Evaluation ebenfalls in den Blick genommenen Phänomene treten dabei global in Erscheinung, wodurch sich in den davon betroffenen Ländern ähnliche Anforderungen an Evaluationen stellen. Auch sind Akteure der Präventions- und Evaluationspraxis international mit ähnlichen Herausforderungen und Schwierigkeiten konfrontiert, wenn es darum geht, Evaluationen zu entwickeln, die Lernprozesse und -chancen bieten, um damit langfristig zu einer besseren Wirkung von Prävention und Interventionen beizutragen (siehe auch Baykal et al. 2021: 5).

Für die internationale Verortung soll eine aktuelle Studie des *Global Public Policy Institute (GPPi)* herangezogen werden. Die Studie hat im Rahmen von *PrEval* Evaluationsstrukturen im Bereich der Extremismusprävention und politischen Bildung in Kanada, Finnland, den Niederlanden und dem Vereinigten Königreich untersucht und daraus Empfehlungen zu Weiterentwicklungspotenzialen für Deutschland formuliert (Baykal et al. 2021). Diese Empfehlungen sollen nun in den Kontext der in den vorangegangenen Kapiteln vorgestellten bundesweiten Befragung gesetzt werden, um daraus Weiterentwicklungspotenziale für Deutschland auch im Bereich der Evaluationskapazitäten abzuleiten.

In der Studie des *GPPi* wurden diese Weiterentwicklungspotenziale anhand von sechs strukturellen „Hebeln“ für solidere, nützlichere Evaluierungspraktiken formuliert (Baykal et al. 2021: 68). Im Folgenden sollen dabei die Hebel *Evaluierungskultur* und *Fähigkeiten zum Evaluierungsmanagement* aufgegriffen werden, an die auch die im Rahmen der Erhebung des *IKG* gewonnenen Daten auf struktureller, Organisations- und Akteurebene anknüpfen können.

Die *GPPi*-Studie beschreibt die Vorbedingungen in den einzelnen Ländern als recht unterschiedlich. Da beispielsweise Großbritannien und die Niederlande in der Vergangenheit vermehrt Anschläge im eigenen Land erleben mussten, herrscht in der Prävention vor allem der Sicherheitsgedanke vor. Finnland verfolgt hingegen einen eher multidisziplinären und gesamtgesellschaftlichen Ansatz. Diese grundsätzlich unterschiedlichen kulturellen Paradigmen im Umgang mit Extremismus wirken sich auch auf struktureller Ebene aus. Dadurch ergeben sich nicht nur Unterschiede in den Evaluationskapazitäten im Vergleich zu Deutschland, sondern auch zwischen den untersuchten Ländern. Nicht zuletzt haben sich zwar überall erste Evaluationsstrukturen im Bereich der Extremismusprävention und politischen Bildung herausbilden können, diese sind aber noch nicht so weit institutionalisiert, dass man bereits von etablierten Evaluationsstrukturen sprechen könnte.

Als ersten wichtigen Hebel zur Festigung von Evaluationsstrukturen und zur Gestaltung einer fruchtbaren Evaluationspraxis benennen Baykal und Kolleg:innen in ihrer Studie daher die Schaffung einer konstruktiven Evaluationskultur: „Eine konstruktive Kultur erzeugt den nötigen Raum, in dem Fehler eingeräumt, neue Ansätze au[s]probiert und manche davon scheitern gesehen werden können und in dem Gelerntes offen ausgetauscht werden kann“ (Baykal et al. 2021: 69–73).

In Ländern, in denen sich bereits eine starke Evaluationskultur aufgrund gemeinsamer „kultureller Normen des Lernens“ (Baykal et al. 2021: 70) herausbilden konnte, steht vor allem der Austausch über Gelerntes im Vordergrund, wodurch Fehler stärker transparent gemacht und Fehlschläge als weniger problematisch betrachtet werden – dies ist zum Beispiel in den Niederlanden der Fall. Diese Lernprozesse führen zu einer stärkeren Professionalisierung und verbessern gleichzeitig die Kapazitäten im Expertisenbereich, beispielsweise dann, wenn Evaluationsansätze regelmäßig in der Präventionspraxis reflektiert sowie in der Evaluation präzisiert und weiterentwickelt werden.

Hierbei übernehmen die Förderinstitutionen eine wichtige Rolle – zumal diese den Großteil der Evaluationen im Bereich der Extremismusprävention und politischen Bildung in den untersuchten Ländern wie auch in Deutschland finanzieren sowie initiieren und somit die strukturelle Grundlage für den Ausbau von Evaluationskapazitäten hinsichtlich einer solchen konstruktiven Evaluationskultur schaffen können. Zudem können nach Baykal et al. allein Förderinstitutionen das hierfür notwendige Vertrauen in die eigene Integrität stärken, also darin, dass Evaluationen z.B. nicht missbraucht werden, um Träger zu sanktionieren, wenn Pilotprojekte nicht die gewünschten Ergebnisse liefern.

Betrachtet man die Forschungslage in Deutschland, so besteht bezüglich einer solchen Evaluations- und Fehlerkultur noch starker Aufholbedarf. Dies spiegelt sich auch in den Stimmungsbildern von Evaluationsakteuren in der bundesweiten Erhebung des *IKG* wider. So zeigt sich in den Antworten ein zum Teil ausgeprägtes Misstrauen gegenüber Förderinstitutionen (allerdings weniger gegenüber Evaluationen oder Evaluierenden):

„Die Idee, dass Träger über Misserfolge berichten, mit dem Vertrauen, dass das anerkannt wird und gesagt wird, wir fördern euch weiter/wieder – das ist nicht existent. [...] Das hat zu einer völligen Verzerrung des Evaluationswesens geführt, und zu einem riesigen Misstrauen. Dabei ist klar, dass in unserem Bereich unmöglich alles funktionieren kann, denn

wir arbeiten mit Leuten, die riesige Probleme haben. Die Entwicklung, die der Bereich der zivilgesellschaftlichen Förderung genommen hat, ist meiner Meinung nach eine, dass es eine Tendenz dazu gibt, misstrauisch zu sein gegenüber externer Evaluation und Eigen-evaluation beschönigend darzustellen – nicht weil man naiv ist, oder die Probleme nicht sieht, sondern weil man Angst davor hat, dass das negative Konsequenzen hat und es nicht gewürdigt wird.“

Ähnliche Aussagen zeigen sich in den Interviewdaten häufiger, insbesondere im Bereich der Selbstevaluation. Soll Selbstevaluation aber der Professionalisierung der Extremismusprävention und Demokratieförderung dienen, dann sind transparente und nachvollziehbare Evaluationsergebnisse notwendig, die ihrerseits gut ausgebildeter Evaluationsstrukturen und -kapazitäten bedürfen. Auch wenn in den Interviews die zum Teil vertrauensvolle Zusammenarbeit mit Mittelgebern positiv herausgestellt wurde, so ist die Skepsis gegenüber den Förderinstitutionen vor allem bei kleineren Trägern teils stärker ausgeprägt. Dadurch werden Lernprozesse erschwert und es kann sich keine vertrauensvolle Evaluations- und Fehlerkultur entwickeln. Demzufolge kann hier nur die Empfehlung der *GPPI*-Studie unterstrichen werden, dass Durchführungsorganisationen vor übertriebenen Konsequenzen bei Fehlschlägen geschützt werden müssen (Baykal et al. 2021: 6, 72) und dieser Schutz ein Teil der Evaluationskultur wird.

Baykal et al. empfehlen für die Prävention in Deutschland daher, dass mehr auf langfristige Förderungen gebaut werden sollte, damit „das organisatorische Überleben von Durchführungsorganisationen nicht ständig auf dem Spiel steht“ (Baykal et al. 2021: 71). Als Negativbeispiel wird in der *GPPI*-Studie in diesem Zusammenhang Großbritannien genannt: Dort hat sich trotz großer Investitionen in Evaluation und wissenschaftliche Forschung eher eine Kultur des Misstrauens und der Geheimhaltung entwickelt, die sich mit der sicherheitszentrierten Herangehensweise der Regierung erklären lässt, welche „wiederum auf eine destruktive Politisierung des gesamten P/CVE-Feldes<sup>10</sup> im UK zurückzuführen [ist]“ (Baykal et al. 2021: 70). Es sei hier bislang nicht gelungen, eine konstruktive Evaluationskultur aufzubauen. Anzeichen für Versicherheitlichung im Bereich der Extremismusprävention sind aber auch in Deutschland wahrnehmbar und müssen ernstgenommen werden (Kiefer 2021: 30). Um Präventions-, aber auch Evaluationsengagement im Bereich der Sicherheit nachvollziehbar zu machen, ist daher grundlegend auf eine größtmögliche Transparenz zu achten.

Nicht unwichtig dürfte in den Niederlanden als Positivbeispiel einer konstruktiven Fehlerkultur auch sein, dass eine eher kleine Gemeinschaft von Evaluationsexpert:innen ein gemeinsames Verständnis bestimmter Ziele teilt und zwischen verschiedenen Rollen in Forschung, Evaluation, Umsetzung und öffentlicher Verwaltung frei hin und her wechselt. Baykal und Kolleg:innen nennen dies einen „Drehtüreffekt“, der für eine offene und konstruktive Lernkultur sowie für ein Gemeinschaftsgefühl sorgt (Baykal et al. 2021: 58).

---

10 P/CVE bedeutet „Preventing and Countering Violent Extremism“, meint also die Prävention und Bekämpfung von gewaltbereitem Extremismus.

Diesem Drehtüreffekt schreiben Baykal und Kolleg:innen einen weiteren positiven Effekt zu, wenn es darum geht, welcher „Markt an Evaluierenden“ am wahrscheinlichsten die beste Evaluierungskapazität für eine Vielzahl verschiedener Evaluierungsbedarfe und -strategien liefert“ (Baykal et al. 2021: 77). Anhand der dargestellten Studienergebnisse des *GPPi* können als Zielvariablen vor allem die Qualifikationsvielfalt, unterschiedliche berufliche Hintergründe und Perspektiven benannt werden. Die institutionelle Praxis des Drehtüreffekts in den Niederlanden scheint „zwischen Regierungsbehörden, regierungsunabhängigen P/CVE-Durchführungsorganisationen und Evaluierungsberatungen insbesondere diesem Ziel einen guten Dienst zu erweisen“ (Baykal et al. 2021: 77). In Kanada gibt es dagegen aufgrund einer weniger differenzierten Zusammensetzung von Akteuren Bedenken wegen Interessenkonflikten, vor allem dann, wenn in Konkurrenz stehende Personen sich an Peer-Evaluationen beteiligen (Baykal et al. 2021: 77). Dem in den Niederlanden positiv diversifizierten Evaluationsmarkt bescheinigen Baykal und Kolleg:innen bessere Voraussetzungen dafür, dass Auftraggebende unabhängige und unparteiliche Evaluierende identifizieren können (Baykal et al. 2021: 77). Dabei spielen aber sicherlich auch wiederum der kleinere Evaluationsmarkt in den Niederlanden und dessen Übersichtlichkeit eine Rolle.

Für die Beurteilung und Weiterentwicklung der Evaluationslandschaft in Deutschland zeigen sich dabei Anknüpfungspunkte. Wie in Kapitel 4 dargestellt, existiert in Deutschland ein breit aufgestellter Evaluationsmarkt (die *PrEval*-Plattform liefert hier zusätzliche Werkzeuge, um diesen anwendungsfallbezogen zu analysieren). Um den Evaluationsmarkt gezielter strukturell zu stärken, könnten Evaluationenaufträge entsprechend den jeweils aktuellen Bedarfen vergeben und die Ressourcenzuleitung gesteuert werden. Dadurch könnte aktuell beispielsweise die Grundlagenforschung im Bereich der Extremismusprävention gestärkt werden – also jene Forschung, die entlang klarer Kriterien einen unabhängigen Blick auf die Prävention richten kann. Dies stellt einen sehr wichtigen Bedarf dar, der in der bundesweiten Befragung des *IKG* immer wieder geäußert wurde. Die Interviewten gaben dabei an, dass in Deutschland generell noch ein Umsetzungsproblem von Forschungserkenntnissen bestünde – der Transfer von der Theorie in die Praxis erfolge dementsprechend nur sehr langsam. Durch schnelle gesellschaftliche Veränderungen sei es zudem schwierig, mit der Forschung den Bedarf in der Praxis zu decken und auch insgesamt sollte die Wissenschaft aktiver auftreten. Gerade in der Extremismusprävention drängt die Zeit zwischen der Identifikation und der Bearbeitung von Phänomenen, was eine schnellere Vermittlung von Forschungswissen erforderlich macht. Zudem fehlten Langzeitstudien über adressierte Zielgruppen. „Wir wissen eigentlich nur, was kommt am Ende des Projektes heraus, aber wir wissen nicht, was passiert eigentlich mit der Zielgruppe dann 10 Jahre später. [...] Nur Forschung kann quasi solche nachgehenden Dinge machen“, so eine Interviewpartnerin im Bereich der Fremdevaluation. Auch hier empfehlen sich auf Dauer angelegte Kooperationen zwischen Prävention und Evaluation und auch hierfür sind wiederum entsprechende finanzielle und zeitliche Ressourcen notwendig.

In Deutschland gibt es bisher auch keine oder nur geringe Evaluationskapazitäten im Bereich der Zusammenführung und Bündelung von Evaluationsergebnissen. Größtenteils erfolgt dies derzeit eher über Forschungsarbeiten. So sollte in Deutschland verstärkt in Kapazitäten für das Management von Evaluationen und in die Verwertung der Ergebnisse investiert werden. Auch in den Interviews der bundesweiten Befragung wurde dieser Bedarf gesehen. Ein Interviewpartner betonte, dass

ein übergeordnetes bzw. staatlich (gefördertes) Evaluationsmanagement hierfür sinnvoll sein könnte – „es kommt aber auf die Ansprache an. Ich hoffe, dass es jemanden gibt, der die Gesamtübersicht behält, aber ich bin in den letzten 20 Jahren nicht überrascht worden mit strategisch, qualitativ guten Entscheidungen.“ Insgesamt wäre es jedoch hilfreich, einen institutionellen Ansprechpartner für Qualitätssicherung und Evaluation zu haben und Evaluation strategischer anzugehen, so eine weitere Interviewpartnerin der Selbstevaluation. Einen Überblick über hierfür notwendige Kapazitätenbedarfe kann die *PrEval*-Plattform liefern. Vor allem bei einem fortlaufendem Mapping könnten Förderinstitutionen beim Erkennen von Kapazitäts- und/oder Nutzungslücken die „Ressourcen und entsprechenden Regelungen ihrer Fördersysteme ausnutzen, um zusätzliche unabhängige und unparteiische Evaluierungskapazitäten aufzubauen“ (Baykal et al. 2021: 77).

## 6. FAZIT UND AUSBLICK

Die bundesweite Befragung und das Mapping der Evaluationskapazitäten zeigen, dass die Kapazitäten in Deutschland im Bereich der Extremismusprävention sowie den Schnittstellen zur Gewaltprävention und politischen Bildung breit aufgestellt sind. Um die Kapazitäten zu stärken und sinnvoll weiterzuentwickeln, sollten aber wichtige Weichenstellungen erfolgen.

Um die vielfältigen Aufgaben, die an Evaluation gestellt werden, erfüllen zu können, sollten Ressourcen gleichermaßen erhöht und gezielter eingesetzt werden. Um Evaluationsansätze und -modelle kontinuierlich testen, kombinieren und weiterentwickeln zu können, sind mehr auf Dauer angelegte Finanzierungsmodelle für Prävention und Evaluation erforderlich. Auch sollten Prävention und Evaluation noch stärker verzahnt werden, um gegenseitige Lernprozesse besser verbinden zu können und einen schnelleren Wissensaustausch zu ermöglichen.

Die Teilnehmenden der bundesweiten Befragung äußerten dabei einen starken Bedarf an Grundlagenforschung im Bereich der Radikalisierung. Dabei können Hochschuleinrichtungen und -institute eine wichtige Rolle übernehmen. Allerdings sind die Kapazitäten hier noch ausbaufähig und bedürfen vor allem einer besseren Personal- und Ressourcenausstattung. Diese könnte unter anderem durch entsprechende Förderungen im Rahmen von öffentlichen Ausschreibungsverfahren erfolgen. Aber auch bei der Entwicklung eines *Bundesinstituts Qualitätssicherung*, wie von der letzten Bundesregierung im Kabinettsausschuss zur Bekämpfung von Rechtsextremismus und Rassismus vorgeschlagen (BMI 2021), könnten entsprechende Ressourcen mitbedacht werden. Ein solches Institut könnte als zentrale Beratungs- und Koordinierungsstelle auftreten und dabei auch den Wissenstransfer von der Wissenschaft in die Praxis vereinfachen. Damit ließen sich Evaluationskapazitäten bündeln sowie effektiver einsetzen und somit auch für deren Entlastung sorgen.

Im „Evaluationsdreieck“ zwischen Wissenschaft, Politik und Praxis (Walkenhorst 2019) bedarf es hierzu immer eines Abgleichs zwischen den verschiedenen Perspektiven, Handlungslogiken, aber vor allem eines gegenseitigen Verständnisses, um ein vertrauensvolles Evaluationsumfeld zu schaffen (siehe auch Baykal et al. 2021: 6). Die Perspektive der Evaluierenden scheint bisher eher weniger Beachtung gefunden zu haben. Evaluierenden wird einerseits oft eine Außenperspektive

zugeordnet, andererseits werden aber Leistungen (z.B. gute Handlungsempfehlungen) erwartet, die nur aus einer Innenperspektive heraus zu erbringen sind. Starke und auf Langfristigkeit angelegte Arbeitsnetzwerke sollten alle am Evaluationsdreieck Beteiligten einschließen, da diese ohnehin in verschiedenen Abhängigkeitsverhältnissen zueinander stehen. Wenn die Prävention mit zu wenigen Finanzmitteln ausgestattet ist, dann wirkt sich das auch auf die Evaluation aus. Wenn der Prävention zu wenig Zeit eingeräumt wird, hat das auch Auswirkungen auf die Zeit, die für begleitende Evaluationen zur Verfügung steht.

Auch wenn gesellschaftliche Entwicklungen ein zum Teil sehr hohes Tempo vorlegen, so benötigt ein methodisches Vorgehen mit ausgewogenem Einbezug aller Stakeholder und Verantwortlichen eine gewisse Zeit – zumeist mehr Zeit als in der Praxis zur Verfügung steht. Deshalb ist es umso wichtiger, dass Evaluation frühzeitig ansetzt und ansetzen kann. Auch hierfür müssen strukturelle Möglichkeiten geschaffen werden und damit wird die Frage nach geeigneten Evaluationskapazitäten umso wichtiger.

- Baykal, Asena/Bressan, Sarah/Friedrich, Julia/Pasquali, Giulia/Rotmann, Philipp/Wagner, Marie 2021: Extremismusprävention evaluieren: Institutionelle Strukturen im internationalen Vergleich, URL: [https://preval.hsfk.de/fileadmin/HSFK\\_Preval/GPPi\\_2021\\_Extremismuspraevention\\_evaluieren\\_Institutionelle\\_Strukturen\\_i....pdf](https://preval.hsfk.de/fileadmin/HSFK_Preval/GPPi_2021_Extremismuspraevention_evaluieren_Institutionelle_Strukturen_i....pdf) (letzter Zugriff: 30.01.2022).
- Becker, Helle/Fereidooni, Karim/Krüger, Thomas/Nordbruch, Götz/Oberle, Monika 2020: Politische Bildung und Primärprävention. Auszug aus einer Fachdebatte, in: Hößl, Stefan E./Jamal, Lobna/Schellenberg, Frank (Hrsg.): Politische Bildung im Kontext von Islam und Islamismus, Bonn, 165–185.
- Beywl, Wolfgang/Kehr, Jochen/Mäder, Susanne/Niestroj, Melanie 2007: Evaluation Schritt für Schritt. Planung von Evaluationen (Hiba-Weiterbildung : 20, Qualifizierung und Beschäftigung in NRW; Bd. 26), Münster.
- BMFSFJ 2021: Programmevaluation, URL: <https://www.demokratie-leben.de/das-programm/programmevaluation> (letzter Zugriff: 30.01.2022).
- BMFSFJ 2020: Grundsätze der Förderung im Handlungsbereich Land im Rahmen des Bundesprogramms „Demokratie leben!“, 10.02.2020, URL: [https://www.demokratie-leben.de/fileadmin/Demokratie-Leben/Downloads\\_Dokumente/Foerdergrundsaeetze/200210\\_grundsaeetze\\_der\\_foerderung\\_land.pdf](https://www.demokratie-leben.de/fileadmin/Demokratie-Leben/Downloads_Dokumente/Foerdergrundsaeetze/200210_grundsaeetze_der_foerderung_land.pdf) (letzter Zugriff: 30.01.2022).
- BMI 2021: Abschlussbericht des Kabinettsausschusses zur Bekämpfung von Rechtsextremismus und Rassismus, URL: [https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/sicherheit/abschlussbericht-kabinettsausschuss-rechtsextremismus.pdf?\\_\\_blob=publicationFile&v=2](https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/sicherheit/abschlussbericht-kabinettsausschuss-rechtsextremismus.pdf?__blob=publicationFile&v=2) (letzter Zugriff: 30.01.2022).
- Brinkmann, Henrik/Kerlen, Christiane/Bilaine, Katharina 2019: Interne Evaluation mit externer Begleitung – Ein Erfahrungsbericht aus dem Non-Profit-Bereich, in: Zeitschrift für Evaluation 18: 1, 29–47.
- DeGEval 2017: Standards für Evaluation. Erste Revision 2016, Mainz, URL: [https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards\\_fuer\\_Evaluation.pdf](https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards_fuer_Evaluation.pdf) (letzter Zugriff: 30.01.2022).
- DeGEval 2008: Empfehlungen für die Aus- und Weiterbildung in der Evaluation. Anforderungsprofile an Evaluatorinnen und Evaluatoren, Mainz, URL: [https://www.degeval.org/fileadmin/Publikationen/Publikationen\\_Homepage/DeGEval\\_Empfehlungen\\_Aus-\\_und\\_Weiterbildung\\_Screen.pdf](https://www.degeval.org/fileadmin/Publikationen/Publikationen_Homepage/DeGEval_Empfehlungen_Aus-_und_Weiterbildung_Screen.pdf) (letzter Zugriff: 30.01.2022).
- DeGEval 2004: Empfehlungen zur Anwendung der Standards für Evaluation im Handlungsfeld der Selbstevaluation, Alfter, URL: [https://www.degeval.org/fileadmin/Publikationen/DeGEval\\_-\\_Empfehlungen\\_Selbstevaluation.pdf](https://www.degeval.org/fileadmin/Publikationen/DeGEval_-_Empfehlungen_Selbstevaluation.pdf) (letzter Zugriff: 30.01.2022).
- Ehnert, Katrin/Johann, Tobias/Mielke, Anne/Rehse, Aline/Walter, Elisa/Zimmermann, Eva/unter Mitarbeit von Clemens Wagner 2021: Handlungsfeld und Handlungslogiken. Wissenschaftliche Begleitung der Modellprojekte im Handlungsfeld Demokratieförderung im Bundesprogramm „Demokratie leben!“ in der Förderphase 2020 bis 2024, München, URL: [https://www.demokratie-leben.de/fileadmin/Demokratie-Leben/Downloads\\_Dokumente/WB-Berichte\\_2.\\_Foerderperiode/Handlungsbereich\\_Modellprojekte/Erster\\_Bericht\\_MP\\_DF\\_2020.pdf](https://www.demokratie-leben.de/fileadmin/Demokratie-Leben/Downloads_Dokumente/WB-Berichte_2._Foerderperiode/Handlungsbereich_Modellprojekte/Erster_Bericht_MP_DF_2020.pdf) (letzter Zugriff: 30.01.2022).
- Freiheit, Manuela/Uhl, Andreas/Zick, Andreas 2021a: Ein systematischer Blick auf die Präventionslandschaft in Deutschland, in: MAPEX-Forschungsverbund (Hrsg.): Radikalisierungsprävention in Deutschland. Mapping und Analyse von Präventions- und Distanzierungsprojekten im Umgang mit islamistischer Radikalisierung, Osnabrück/Bielefeld, 49–81.

- Freiheit, Manuela/Uhl, Andreas/Zick, Andreas 2021b: Die Präventionslandschaft in Deutschland, in: Bundesamt für Migration und Flüchtlinge (Hrsg.): Deradikalisierungs- und Distanzierungsarbeit. Begleitbuch zum Qualifizierungslehrgang (Umfeld-)Beratung im Phänomenbereich islamistisch begründeter Extremismus. Beiträge zu Migration und Integration, Nürnberg 9, 124–130.
- Kiefer, Michael 2021: Radikalisierungsprävention in Deutschland – Ein Problemaufriss, in: MAPEX-Forschungsverbund (Hrsg.): Radikalisierungsprävention in Deutschland. Mapping und Analyse von Präventions- und Distanzierungsprojekten im Umgang mit islamistischer Radikalisierung, Osnabrück/Bielefeld, 29–48.
- Koynova, Svetla/Mönig, Alina/Quent, Matthias/Ohlenforst, Vivienne 2022: Monitoring, Evaluation und Lernen: Erfahrungen und Bedarfe der Fachpraxis in der Prävention von Rechtsextremismus und Islamismus, PRIF Report 7/2022, Frankfurt am Main.
- Labin, Susan N./Duffy, Jennifer L./Meyers, Duncan C./Wandersman, Abraham/Lesesne, Catherine A. 2012: A Research Synthesis of the Evaluation Capacity Building Literature, in: American Journal of Evaluation 33: 3, 307–338.
- Lützing, Saskia/Gruber, Florian/Hedayat, Ali 2020: Extremismuspräventionslandschaft – eine Bestandsaufnahme präventiver Angebote in Deutschland sowie ausgewählter Präventionsstrategien aus dem europäischen Ausland, in: Ben Slama, Brahim/Kemmesies, Uwe (Hrsg.): Handbuch Extremismusprävention. Gesamtgesellschaftlich. Phänomenübergreifend, Wiesbaden, 597–626.
- Milbradt, Björn/Schau, Katja/Greuel, Frank 2019: (Sozial-)pädagogische Praxis im Handlungsfeld Radikalisierungsprävention – Handlungslogik, Präventionsstufen und Ansätze, in: Heinzemann, Claudia/Marks, Erich (Hrsg.): Prävention & Demokratieförderung. Gutachterliche Stellungnahmen zum 24. Deutschen Präventionstag, Bonn, 141–179.
- Möller, Kurt/Buschbom, Jan/Pfeiffer, Thomas 2020: Zur Evaluation von Praxisansätzen der Extremismusprävention, in: Ben Slama, Brahim/Kemmesies, Uwe (Hrsg.): Handbuch Extremismusprävention. Gesamtgesellschaftlich. Phänomenübergreifend, Wiesbaden, 389–419.
- Müller-Kohlenberg, Hildegard 2006: Evaluation und Selbstevaluation in der Sozialen Arbeit, in: Galuske, Michael/Thole, Werner (Hrsg.): Vom Fall zum Management: Neue Methoden der Sozialen Arbeit, Wiesbaden, 71–81.
- Nehlsen, Inga/Biene, Janusz/Coester, Marc/Greuel, Frank/Milbradt, Björn/Armborst, Andreas 2020: Evident and Effective? The Challenges, Potentials and Limitations of Evaluation Research on Preventing Violent Extremism, in: International Journal of Conflict and Violence 14: 2, 1–20.
- OECD 2021: Developing Evaluation Capacities, URL: <https://www.oecd.org/dac/evaluation/evaluatingcapacitydevelopment.htm> (letzter Zugriff: 30.01.2022).
- Patton, Michael Q. 2008: Utilization-Focused Evaluation, 4. Ausgabe., Thousand Oaks, CA.
- Silvestrini, Stefan 2016: Ex Ante Evaluation as a Precondition for Evaluating Impact, in: Stockmann, Reinhard/Meyer, Wolfgang (Hrsg.): The Future of Evaluation, London, 269–281.
- Stockdill, Stacey/Baizerman, Michael/Compton, Donald W. 2002: Toward a Definition of the ECB Process: A Conversation with the ECB Literature, in: New Directions for Evaluation 93, 7–26.
- Stockmann, Reinhard/Meyer, Wolfgang/Taube, Lena 2020: The Institutionalisation of Evaluation in Europe, Cham.

Stockmann, Reinhard 2018: Der Evaluation Capacity Development Ansatz des Centrums für Evaluation (CEval). (CEval-Arbeitspapier, 21), Saarbrücken, URL: <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-60720-8> (letzter Zugriff: 30.01.2022).

Stockmann, Reinhard 2010: Rolle der Evaluation in der Gesellschaft, URL: <https://publikationen.sulb.uni-saarland.de/handle/20.500.11880/25680> (letzter Zugriff: 30.01.2022).

Stockmann, Reinhard 2008: Zur gesellschaftlichen Bedeutung von Evaluation. (CEval-Arbeitspapier, 15), Saarbrücken, URL: [https://www.ceval.de/modx/fileadmin/user\\_upload/PDFs/workpaper15.pdf](https://www.ceval.de/modx/fileadmin/user_upload/PDFs/workpaper15.pdf) (letzter Zugriff: 30.01.2022).

Walkenhorst, Dennis 2019: Das „Erwartungsdreieck Evaluation“: Eine Praxisperspektive, 20.03.2019, URL: <https://www.bpb.de/politik/extremismus/radikalisierungspraevention/287931/das-erwartungsdreieck-evaluation-eine-praxisperspektive> (letzter Zugriff: 30.01.2022).

## PRIF REPORT

Die PRIF Reports analysieren Hintergründe politischer Ereignisse und Entwicklungen und präsentieren wissenschaftliche Forschungsergebnisse in Deutsch oder Englisch.

Junk, Julian/Koynova, Svetla/Ohlenforst, Vivienne/Ruf, Maximilian/Scheu, Lea (2022): Evaluation und Qualitätssicherung in der Kooperation sicherheitsbehördlicher und zivilgesellschaftlicher Akteure in der sekundären und tertiären Extremismusprävention, PRIF Report 8/2022, Frankfurt/M.

Koynova, Svetla/Mönig, Alina/Quent, Matthias/Ohlenforst, Vivienne (2022): Monitoring, Evaluation und Lernen: Erfahrungen und Bedarfe der Fachpraxis in der Prävention von Rechtsextremismus und Islamismus, PRIF Report 7/2022, Frankfurt/M.



[www.hsfk.de/PRIF-Reports](http://www.hsfk.de/PRIF-Reports)

[www.hsfk.de/HSFK-Reports](http://www.hsfk.de/HSFK-Reports)

## PRIF SPOTLIGHT

Die PRIF Spotlights diskutieren aktuelle politische und gesellschaftliche Themen.

Pfeifer, Hanna/Al Malla, Houssein/Weiß, David (2022): "Enlist Now!" – Or Don't? Why we should be concerned about foreign fighting in Ukraine, PRIF Spotlight 5/2022, Frankfurt/M.

Wolff, Jonas (2022): Deutsche Demokratieförderpolitik. Ampelkoalition auf Kurswechsel?, PRIF Spotlight 4/2022, Frankfurt/M..

## PRIF SPOTLIGHT

### DIE RICHTIGE FRAU AM RICHTIGEN ORT

Die Bedeutung des Geschlechter-Friedensplans 2016 an Fatma Mghariah

Am 16. und 17. April 2022 fand die PRIF-Tagung 'Die richtige Frau am richtigen Ort' statt. In der Tagung wurden die Erfahrungen von Frauen in Konfliktsituationen diskutiert und die Bedeutung des Geschlechter-Friedensplans 2016 an Fatma Mghariah hervorgehoben.



Die PRIF-Tagung 'Die richtige Frau am richtigen Ort' wurde am 16. und 17. April 2022 in Frankfurt am Main durchgeführt. Die Tagung wurde von der PRIF und dem HSFK organisiert und wurde von der PRIF und dem HSFK unterstützt.

[www.hsfk.de/PRIF-Spotlights](http://www.hsfk.de/PRIF-Spotlights)

## PRIF BLOG

Auf dem PRIF Blog erscheinen Beiträge zu aktuellen politischen Fragen und Debatten der Friedens- und Konfliktforschung. Die Blogbeiträge erscheinen in loser Folge in Deutsch oder Englisch.

## PRIF BLOG

PRIF RESEARCH INSTITUTE FRANKFURT | LEIBNIZ-INSTITUT FÜR FRIEDENS- UND KONFLIKTFORSCHUNG

<https://blog.prif.org/>

PRIF Reports und PRIF Spotlights sind Open-Access-Publikationen und können kostenlos auf [www.hsfk.de](http://www.hsfk.de) heruntergeladen werden. Sie möchten die digitalen Ausgaben abonnieren? Bitte wenden Sie sich an: [publikationen@hsfk.de](mailto:publikationen@hsfk.de).

[www.facebook.com/HSFK.PRIF](https://www.facebook.com/HSFK.PRIF)

[www.twitter.com/HSFK\\_PRIF](https://www.twitter.com/HSFK_PRIF)

<https://blog.prif.org/>

Gefördert durch:



Bundesministerium  
des Innern, für Bau  
und Heimat

aufgrund eines Beschlusses  
des Deutschen Bundestages

**PrEval**

Evaluationsdesigns für Präventionsmaßnahmen –  
Multimethodische Ansätze zur Wirkungsermittlung und  
Qualitätssicherung in der Extremismusprävention sowie den  
Schnittstellen zur Gewaltprävention und politischen Bildung

ANDREAS UHL // MANUELA FREIHEIT //  
BENJAMIN ZEIBIG // ANDREAS ZICK

## EVALUATIONSKAPAZITÄTEN IM BEREICH DER EXTREMISMUSPRÄVENTION UND DER POLITISCHEN BILDUNG IN DEUTSCHLAND

Evaluationen erfüllen vielfältige gesellschaftliche Funktionen und werden auch in der deutschen Extremismusprävention immer wichtiger. Wie gewinnbringend und effektiv sie sind, hängt jedoch wesentlich von bestehenden Evaluationskapazitäten ab.

Mithilfe einer bundesweiten Telefonbefragung unter 104 Evaluationsakteuren haben die Autor:innen die Evaluationskultur sowie vorhandene Kapazitäten und Ressourcen im Bereich der Fremd- und Selbstevaluation erhoben. Der vorliegende Report analysiert die gesammelten Daten und bietet einen systematischen Überblick über bestehende Evaluationskapazitäten in Deutschland sowie die Weiterentwicklungspotenziale für eine effektive und konstruktive Evaluationspraxis im Bereich der Extremismusprävention und politischen Bildung.

ISBN 978-3-946459-78-1