

Impulse für evidenzbasierte Politikgestaltung

**Praxisnahe Tipps und inspirierende Fallbeispiele für
wirksamere Entscheidungen**

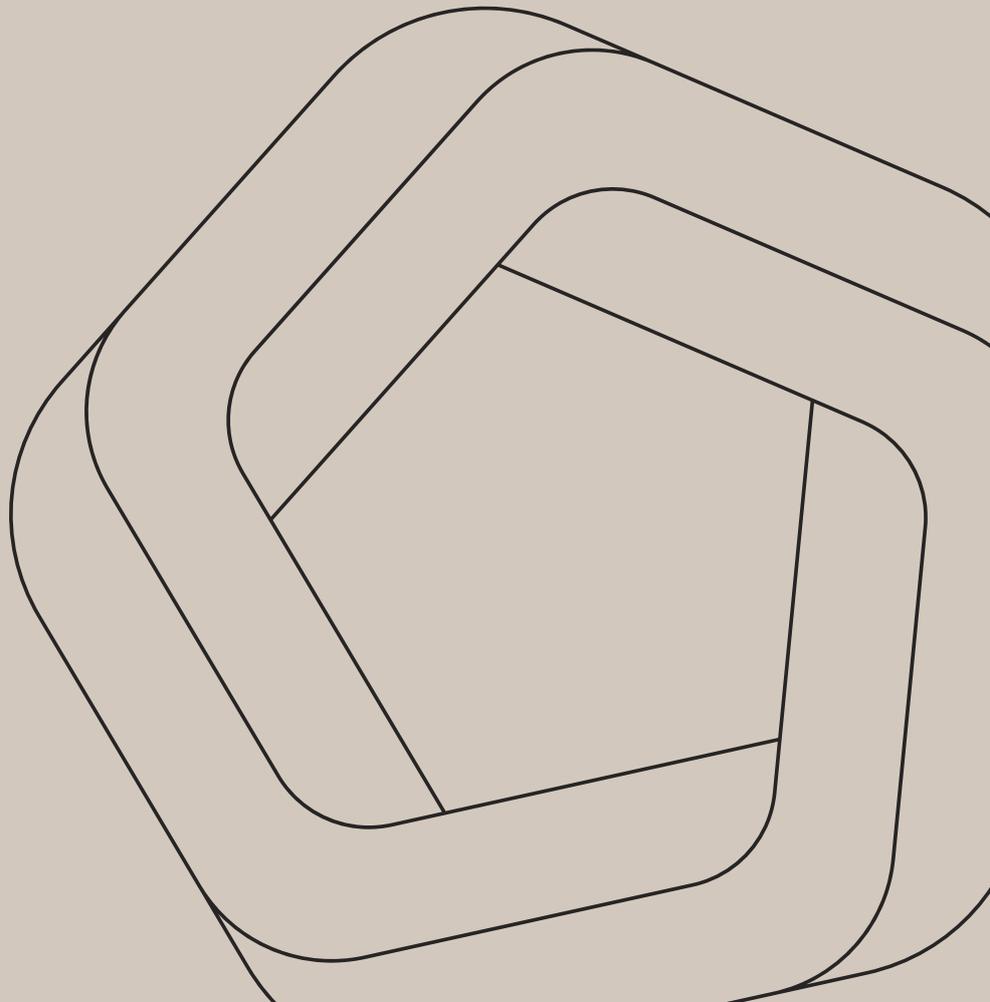
Handreichung

gefördert von der Robert-Bosch-Stiftung

Von Elke Himmelsbach

Dr Sophia McDonnell

2025



Inhalt

Zusammenfassung	3
Warum evidenzbasierte Entscheidungsfindung?	4
Was ist gute Evidenz?	6
Die Verwendung von Evidenz im Politikzyklus	8
DIAGNOSE: Ein Problem und seine Ursachen verstehen	9
DESIGN: Entwerfen und Auswählen von Lösungen	13
IMPLEMENT: Feinabstimmung der Implementierung einer Lösung	18
EVALUATE: Bewertung von Leistungen, Wirkungen und langfristigen Effekten einer Lösung	22
Literaturhinweise	25

Zusammenfassung

Impulse für evidenzbasierte Politikgestaltung – Warum und wie evidenzbasierte Entscheidungen gelingen können

Diese Handreichung zeigt praxisnah, wie wissenschaftliche und erfahrungsbasierte Evidenz entlang des gesamten Politikzyklus systematisch genutzt werden kann – von der Problemdefinition über die Entwicklung und Umsetzung von Lösungen bis hin zur Evaluation. Sie ist mit inspirierenden Fallbeispielen aus Deutschland, Europa und Australien angereichert und richtet sich an politische Entscheidungsträger*innen, die fundiertere, wirksamere und gesellschaftlich akzeptierte Politik gestalten wollen.

Relevanz für die politische Praxis:

Viele politische Entscheidungen erfolgen heute unter Zeitdruck und auf Basis etablierter Routinen, persönlicher Einschätzungen oder politischer Zwänge. Dabei ist klar: Mehr und bessere Nutzung von Evidenz führt nachweislich zu effizienteren Maßnahmen, größerer Bürgernähe und höherer Wirkung – gerade angesichts zunehmender Komplexität und Polarisierung politischer Themen.

Hauptinhalte und Nutzen:

- **Was ist „gute Evidenz“?**

Differenzierung zwischen wissenschaftlicher (quantitativer/qualitativer) Evidenz und Erfahrungsevidenz. Beide Evidenztypen sind relevant – idealerweise im Zusammenspiel.

- **Evidenz im Politikzyklus:**

- » Agenda-Setting: Priorisierung mit Blick auf Dringlichkeit, Datenlücken und politische Machbarkeit
- » Diagnose: Ursachenanalysen mit verhaltenswissenschaftlicher Perspektive
- » Design: Entwicklung, Adaption und Testen wirksamer Lösungsansätze
- » Implementierung: Überwindung von Widerständen, Sicherung der Einbindung von Stakeholdern sowie Skalierbarkeit
- » Evaluation: Messung von Wirkung, Wirtschaftlichkeit, Langzeitfolgen und Lernschleifen

- **Fallstudien mit Praxisbezug:**

Beispiele u.a. zur Bekämpfung von Gewalt gegen Frauen in Australien, zur Helmnutzung in den Niederlanden, zur Lehrerbildung in Deutschland oder zur Recyclingpolitik in Frankreich zeigen: Evidenz kann nicht nur politisches Handeln verbessern, sondern auch gesellschaftliche Akzeptanz stärken und Konflikte entschärfen.

Wichtige Empfehlungen für Entscheidungsträger*innen:

- Gute Evidenz ist vertrauenswürdig, objektiv, relevant, aktuell und methodisch solide.
- Erfolgreiche evidenzbasierte Politik braucht: niedrighschwellige Instrumente, klare Prozesse, interdisziplinäre Zusammenarbeit und Datenkompetenz.
- Auch bei hohem politischen Druck hilft Evidenz, legitime, transparente Entscheidungen zu treffen und Vertrauen aufzubauen.

Fazit:

Diese Handreichung liefert konkrete Handlungsimpulse, um evidenzbasierte Politik in die tägliche Arbeitspraxis zu integrieren. Sie zeigt nicht nur dass, sondern auch wie es geht. Ein unverzichtbarer Werkzeugkasten für alle, die öffentliche Wirkung steigern und politisches Vertrauen stärken wollen.

Warum evidenzbasierte Entscheidungsfindung?

Es besteht ein wachsender Bedarf an evidenzbasierter Politikgestaltung. Wenn politische Antworten, Gesetze und die Umsetzung politischer Programme nicht mit der wahrgenommenen Realität vor Ort übereinstimmen, nimmt die Enttäuschung und Unzufriedenheit bei Bürgerinnen und Bürgern sowie Unternehmen zu (Ludewig in NKR 2019).

Die Forschung zur evidenzbasierten Politikgestaltung zeigt (Sager et al. 2021), dass die Integration von mehr Evidenz in politische Entscheidungen die Wirksamkeit und Effizienz von Vorhaben verbessert und dazu beitragen kann, das Vertrauen der Bürgerinnen und Bürger zurückzugewinnen.

Eine aktuelle Umfrage der Nationalen Akademie der Wissenschaften Leopoldina (Riphahn und Schnitzer 2022) ergab jedoch, dass die politischen Entscheidungsträger im Bundestag wissenschaftliche Erkenntnisse nicht in dem Maße verwenden, wie sie es wünschen und wie es die OECD (2021) und andere nationale Audits (NKR 2019) nahelegen.

Hindernisse für die Verwendung von Evidenz:

- **Gewohnheit oder Tradition**, persönliche Expertise anstelle von Daten zu nutzen (Hadorn et al. 2022); die Einbeziehung von Sozialpartnern und Sachverständigen ist gut etabliert, aber es gibt keine Mechanismen für andere Zielgruppen wie die Einbeziehung der Bürgerinnen und Bürger (OECD 2021)
- **Mangelnde Zugänglichkeit**, d.h. die Evidenzbereitstellung wird von langjährig institutionalisierten Wissenstransfers mit veralteten Verfahren dominiert (NKR 2019, Hadorn et al 2022)
- **Mangel an guter Kommunikation** von wissenschaftlicher Evidenz zwischen Forschenden und Entscheidungsträgern, z. B. Mangel an kurzen und verständlichen Zusammenfassungen, Visualisierungen, transparenten Methoden, unvoreingenommenen Ansätzen, klarer Differenzierung von Empfehlungen, die in den Kontext des Entscheidungsprozesses passen (Seidel et al. 2021)
- **Zeitmangel**, d.h. politische Entscheidungsträger werden oft durch schnellere Medienzeiten unter Druck gesetzt, sich schnell auf scheinbar offensichtliche Lösungen festzulegen, die vielleicht noch gar nicht ausgereift sind, und dies führt zu der Angewohnheit, Lösungen zu schnell und zu detailliert zu definieren (NKR 2019)

Politische Entscheidungsträger haben grundsätzlich Zugang zu Evidenz, aber viele können sie nicht angemessen nutzen. In der Praxis greifen politische Entscheidungsträger auf moralische Werte, persönliche Erfahrungen und politische Notwendigkeiten neben Evidenz zurück, um Entscheidungen zu treffen, sodass sie eine eingeschränkte oder "begrenzte Rationalität" anwenden. Trotz der häufigen Betonung der Bedeutung von Evidenz bleibt die angemessene Verwendung von Evidenz für die Politikgestaltung eine Herausforderung.

Daher besteht ein Bedarf an besseren Werkzeugkästen, um es politischen Entscheidungsträgern einfach und praktikabel zu machen, Evidenz in ihrer täglichen Arbeit in Deutschland zu nutzen. In einem ersten Schritt soll diese Handreichung Anregungen dazu geben, wie diese Hindernisse überwunden und Evidenz in verschiedenen Phasen des Politikzyklus genutzt werden kann.

Die Vorteile der Verwendung von Evidenz:



**Optimierung der
Ressourcenallokation**



**Verbesserung der Erbringung
öffentlicher Dienstleistungen**



**Erzielung von besseren Ergebnissen
für die Gesellschaft**

Was ist gute Evidenz?

Wenn wir über Evidenz sprechen, empfehlen wir, die folgenden zwei Arten von Evidenz zu unterscheiden (WHO 2021):

Wissenschaftliche oder Forschungsevidenz

wird systematisch durch einen formalen, standardisierten Forschungsprozess nach bestimmten methodischen Standards generiert.

Dies kann Primärforschung, Synthese bestehender Forschung (Sekundärforschung) und andere Produkte umfassen, die auf Evidenz basieren, wie z.B. Leitfäden (Tertiärforschung). Wissenschaftliche Evidenz kann numerisch (quantitativ) oder deskriptiv (qualitativ) sein, wobei eine Reihe von Methoden zum Einsatz kommt.

Erfahrungsevidenz

stellt eher informelles Wissen dar, das auf den Meinungen und Erfahrungen von Personen basiert, die in dem zu analysierenden Themenfeld arbeiten oder leben.

Diese Art von Evidenz wird häufig in den Medien, in Expertenberichten, persönlichen Anekdoten, qualitativen Interviews, Gruppendiskussionen oder deliberativen Dialogen geäußert.

Beide Arten von Evidenz können wertvoll sein und unterschiedliche Forschungsfragen beantworten. Ihre Beziehung ist komplementär. Zum Beispiel kann die Erfahrungsforschung einen Bereich hervorheben, auf den sich die wissenschaftliche Forschung konzentrieren sollte, oder sie kann die bereits durchgeführte wissenschaftliche Forschung validieren oder in Frage stellen.

Typische Forschungsfragen, die durch den jeweiligen Evidenztyp beantwortet werden können (Superu 2018), umfassen:

Wissenschaftliche oder Forschungsevidenz

- Bestätigen oder beschreiben die Daten die ursprüngliche Hypothese, die Zielgruppe und/oder die Treiber von Veränderungen?
- Gibt es Belege für die Wirksamkeit von Maßnahmen bei der Lösung eines Problems?
- Gibt es gut konzipierte Studien oder Auswertungen, um die Gelingensbedingungen oder die gewünschten Wirkungen zu validieren?
- Welche positiven, negativen oder unbeabsichtigten Auswirkungen hatte ein Programm oder eine Maßnahme auf Verhaltensänderungen oder andere Ziele?

Erfahrungsevidenz

- Wie spiegelt der politische Ansatz die Erfahrungen und das Wissen der Menschen wider, denen er zugute kommen soll?
- Was können die früheren Erfahrungen einer Zielgruppe über die Akzeptanz, Bedeutung und Angemessenheit der Politik oder des Programms für sie aussagen?
- Was können die Erfahrungen und das Wissen der Stakeholder darüber aussagen, was bisher bei der Zielgruppe funktioniert oder nicht funktioniert hat?
- Welche gemeinsamen Ziele haben die verschiedenen Zielgruppen und Interessensvertretungen in Bezug auf die Politikmaßnahme?

- Gibt es eine Anleitung zur Implementierung? Was sagt die Evidenz über die Ressourcen, Prozesse und Kapazitäten aus, die für die erfolgreiche Umsetzung einer Maßnahme erforderlich sind?
- Was sind die Merkmale der Menschen, die von der Maßnahme profitieren? Wie und in welchem Ausmaß sind die Menschen von dem Thema, Problem oder der Chance betroffen?

- Was sagen die Erkenntnisse darüber aus, wie gut die Erfahrungen und Wünsche der Zielgruppe mit den Zielen der Politikmaßnahme übereinstimmen?
- Welche Erfahrungen wurden aus Evaluationen oder Berichten ähnlicher Initiativen gesammelt? Wie könnten sie genutzt werden, um den Ansatz zu stärken?

Wie kann man gute Evidenz erkennen?

Bei der Überlegung, auf welche Evidenz Entscheidungen gestützt werden sollen, ist es wichtig, die Robustheit und Qualität der verfügbaren Evidenz zu bewerten. Es gibt zahlreiche Leitfäden von Forschungsvereinigungen und -institutionen, die einen detaillierten Überblick über Qualitätsstandards für die Evidenzproduktion geben. In Deutschland bietet die Gesellschaft für Evaluation e.V. (DeGEval) einen guten Überblick über die Standards für Evaluationen und der Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. (<https://www.adm-ev.de/standards-richtlinien/>) stellt eine nützliche Checkliste für Auftraggeber bereit, die Umfragen in Auftrag geben.

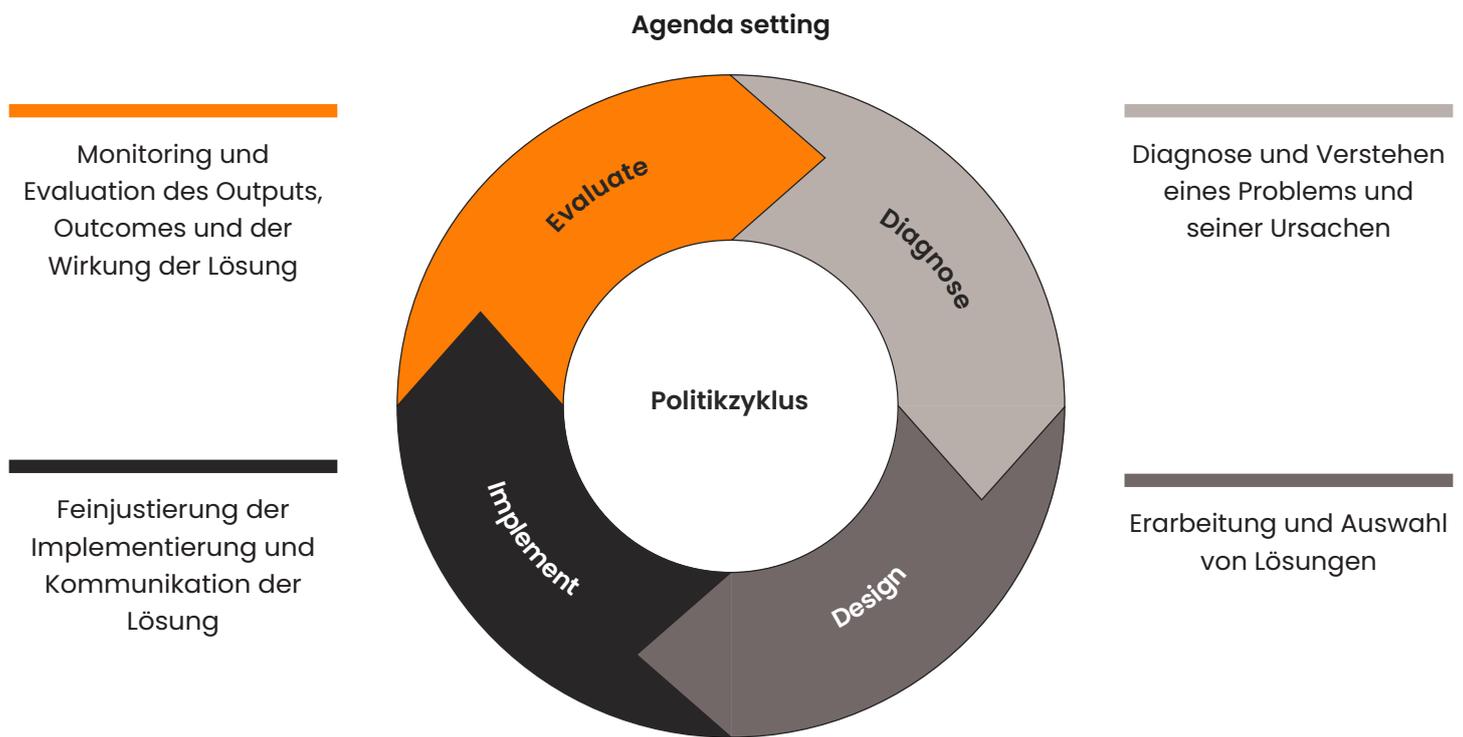
In der Praxis ist es jedoch sinnvoller, sich von prägnanten und klaren Rahmenwerken leiten zu lassen, die Schlüsseldimensionen guter Evidenz hervorheben, wie z. B. das folgende Rahmenwerk aus der evidenzbasierten Politikpraxis im Vereinigten Königreich (Houses of Parliament 2017, Nesta 2022).

Bei der Bewertung von Evidenz aus einer Vielzahl von Quellen und Methoden ist es hilfreich sicherzustellen, dass sie die folgenden Kriterien erfüllt:

VERTRAUENSWÜRDIG	Es ist wichtig zu beurteilen, woher die Informationen stammen und wer sie erstellt hat. Sind die Qualifikationen der Autorinnen und Autoren glaubwürdig und ist ihre Vorgehensweise transparent?
OBJEKTIV	Um das Ziel hinter jedem Forschungsvorhaben und die möglichen Verzerrungen, die auftreten können, zu verstehen, sollten Sie den Zweck der Studie und die wahrscheinlichen Motivationen dahinter berücksichtigen. In diesem Zusammenhang ist es oft sinnvoll zu prüfen, wer die Forschung finanziert hat.
RELEVANT	Nicht jede Forschung ist für Ihre Frage oder Situation nützlich. Sie sollten daher prüfen, ob der Kontext, in dem die Forschung stattgefunden hat, mit Ihrem eigenen vergleichbar ist, ob alle relevanten Perspektiven einbezogen wurden und für welche Zielgruppe die Forschung ursprünglich gedacht war.
AKTUELL	Wenn Sie Ihre Entscheidung auf Evidenz stützen, ist es wichtig, dass diese Daten nicht veraltet und immer noch gültig sind. Sie sollten daher prüfen, ob die Informationen, die Sie haben, aktuell sind und ob in der Zwischenzeit etwas Wichtiges passiert ist, das die Ergebnisse beeinflussen könnte.
GENAU	Um die Qualität einer Methode zu beurteilen, sollten Sie sicherstellen, dass die Methodik detailliert beschrieben und angemessen ist. Hilfreich ist auch zu prüfen, ob ähnliche Studien durchgeführt wurden oder ob die Studie wiederholt wurde. In der Regel ist ein Peer-Review oder eine Validierung durch andere Expertinnen und Experten auf diesem Gebiet ein gutes Zeichen für starke Evidenz.

Die Verwendung von Evidenz im Politikzyklus

Evidenz kann Sie und Ihre politischen Maßnahmen effektiver und leistungsfähiger machen. Dieses Kapitel bietet Ihnen Ideen und inspirierende Fallstudien, die zeigen, wie Evidenz Ihre politische Entscheidungsfindung in verschiedenen Phasen des Politikzyklus unterstützen kann.



Die Entwicklung einer politischen Maßnahme beginnt in der Regel mit einem Agenda-Setting-Prozess, in dem die Themen ausgewählt werden, die als relevant erachtet werden. In der Diagnosephase geht es darum, das Problem und seine Ursachen zu verstehen und zu beschreiben.

Daran schließt sich die Designphase an, in der Lösungen für den nächsten Schritt, die Implementierungsphase, entwickelt und ausgewählt werden. Während der Implementierungsphase werden in der Regel eine Reihe von Entscheidungen getroffen, um die Kommunikation und die Prozesse zu optimieren.

Zu guter Letzt sollte die implementierte Lösung überwacht und evaluiert werden, um festzustellen, ob sie das ursprüngliche Problem gelöst und die gewünschten Ziele erreicht hat. Die Erkenntnisse aus der Evaluation fließen dann in das nächste Agenda-Setting ein, und zwar nicht nur für das gleiche Problem (wenn es nicht gelöst wurde), sondern auch für ähnliche Probleme, auf die sich die Erkenntnisse übertragen lassen.

DIAGNOSE: Ein Problem und seine Ursachen verstehen

Der Politikzyklus beginnt mit der Identifizierung und Formulierung eines Problems, das Aufmerksamkeit erfordert. Im Idealfall sollte diesem Schritt ein formaler Priorisierungsprozess vorausgehen oder ihn begleiten, der Antworten auf folgende Fragen umfasst:

AGENDA-SETTING Sollten wir dieses Problem angehen?	Wie groß ist das Problem? <ul style="list-style-type: none">> Welche Indikatoren eignen sich für die Messung? Wird es schlimmer oder ist es hier größer als anderswo? <ul style="list-style-type: none">> Was sind vergleichbare Benchmarks? Wie lange wird eine Priorisierung des Themas in einer sich ständig verändernden Welt Bestand haben? <ul style="list-style-type: none">> Wann sollte man welche Arten von Evidenz sammeln?> Wann ist die Evidenz veraltet?
---	--

Die Antworten auf die obigen Fragen führen zur Beschreibung des Problems, die je nach Perspektive unterschiedlich aussehen kann. Das Verständnis des Problems sollte daher durch Forschungsergebnisse, den Entscheidungskontext, ethische Erwägungen und Diskussionen mit wichtigen Stakeholdern untermauert werden.

KONTEXT & DIAGNOSE Wie können wir das Problem und seine Ursachen beschreiben?	Was sind die relevanten Kontextszenarien und Lebenssituationen, in denen das Problem auftritt? Wie beschreiben und erleben verschiedene Menschen das Problem und seine Ursachen? <ul style="list-style-type: none">> Wer sind die relevanten Akteure? Wie können die Ursachen des Problems identifiziert werden? <ul style="list-style-type: none">> ... vor allem, wenn sie nicht offensichtlich sind
--	--

Die Fragen, die in der Diagnosephase am schwierigsten zu beantworten sind, sind die jeweils letzteren, weshalb im Folgenden mögliche Lösungen anhand von zwei Fallstudien vorgestellt werden.

Das erste Beispiel verdeutlicht die zeitlichen Herausforderungen in der Agenda-Setting-Phase, die bei Wahlkampfentscheidungen der politischen Parteien auftreten. Es basiert auf den Erfahrungen von Verian in Deutschland.



Wie lange wird eine Priorisierung des Themas in einer sich ständig verändernden Welt Bestand haben? Wann sollte man welche Art von Evidenz sammeln? Wann ist Evidenz veraltet?

Vor den Wahlen müssen politische Parteien Entscheidungen über die Kandidaten und die Themen treffen, auf die sie ihren Wahlkampf stützen wollen. Die Parteigremien entscheiden sich in der Regel für ein evidenzbasiertes Briefing, bevor sie ihre Entscheidungen treffen und die Kommunikationsagentur briefen.

Die größte Herausforderung ist das Timing: Wann wird welche Art von Evidenz gesammelt?

Politische Stimmungen können sehr wechselhaft sein. Ein früher Start ermöglicht ein frühes Briefing mit genügend Zeit, um eine gute Kampagne zu entwickeln. Es birgt aber auch das Risiko, dass viele Dinge passieren können, die die Wahrnehmung und die Wirkung einer Kampagne verändern. **Die Lösung besteht darin, strategische und taktische Ziele zu trennen.**

Typische strategische Forschungsfragen sind:

- Wer sind die zugänglichen Wählersegmente unter den Wahlberechtigten?
- Welche Themen und Narrative mobilisieren potenzielle Wähler*innen und welche Themen demobilisieren politische Gegner?
- Wo ist unsere Partei verwundbar?

Typische taktische Forschungsfragen sind:

- Was ist die zentrale Wahlkampfbotschaft in der letzten Wahlkampfphase, die das beste Potenzial zur Stimmenmaximierung hat?
- Wie geht man mit Briefwähler*innen um, die bis zu sechs Wochen vor dem Wahltermin wählen dürfen?

Evidenz für strategische Fragen wird in der Regel 6 bis 12 Monate vor dem Wahltag gesammelt. Evidenz für taktische Fragen wird etwa 4 bis 12 Wochen vor dem Wahltag gesammelt. **Wenn jedoch unerwartete Ereignisse eintreten** (z. B. ein neuer Konflikt oder ein*e neue*r Gegenkandidat*in), **ist es ratsam, die Strategie auf der Grundlage neuer Erkenntnisse zu aktualisieren, die nach solchen bedeutenden Ereignissen gesammelt wurden.**

Themenunabhängige Empfehlungen für die Agenda-Setting-Phase:

1. Unterteilen Sie die Themen in strategisch/langfristig und taktisch/kurzfristig.
2. Wählen Sie zeitnahe Evidenz, wenn Sie zwischen verschiedenen Problemen, Personen usw. priorisieren.
3. Prüfen Sie, ob die vorhandene Evidenz von wichtigen Ereignissen oder Trends überholt wird.

Das zweite Beispiel verdeutlicht, wie wichtig es ist, ganzheitliche Diagnoseinstrumente einzusetzen, die mehrere Perspektiven einbeziehen, um die Ursachen langfristiger sozialer Probleme zu identifizieren. Nur so ist es möglich, Hebel für veränderte Trends zu identifizieren. Es basiert auf den Erfahrungen von Verian in Australien.



Wie lassen sich Ursachen für ein Problem identifizieren, insbesondere wenn sie nicht offensichtlich sind?

Gewalt gegen Frauen hat in Australien eine lange Tradition. Zwei von fünf australischen Frauen hatten nach dem 15. Lebensjahr Gewalt durch einen Mann erlebt. Jede Woche wurden ein oder zwei Frauen von einem aktuellen oder ehemaligen Partner getötet. Die Folgen von Gewalt gegen Frauen kosteten die australische Wirtschaft jährlich 13,6 Milliarden AUD. Die australische Regierung hat die Verringerung der Gewalt gegen Frauen im Jahr 2015 zu einer nationalen Priorität der öffentlichen Gesundheit gemacht.

Es gab zwar umfangreiche Forschungsarbeiten zu diesem Thema, aber die überwiegende Mehrheit konzentrierte sich darauf, das Ausmaß des Problems und die rationalen, bewussten Einstellungen zu identifizieren, die es beeinflussen können.

Es fehlte jedoch das Verständnis für den Kern des Problems, seinen Ursprung und daher auch Ideen, was getan werden könnte, um Gewalt zu reduzieren.

Neue Forschungsergebnisse zeigten, dass soziale Normen und unbewusste Überzeugungen bei Erwachsenen und Jugendlichen Geschlechterungleichheit, Respektlosigkeit und aggressives Verhalten fördern. Zum Beispiel gab es eine Tendenz zur Täter-Opfer-Umkehr, wenn Beispiele für aggressives oder respektloses Verhalten zwischen Männern und Frauen genannt wurden („sie muss etwas getan haben“). Außerdem wurde häufig Empathie mit den männlichen Tätern gezeigt („Es ist sehr schwer, ein junger Mann zu sein, er mag sie wahrscheinlich nur“) oder das Verhalten verhamlost („es ist nicht so schlimm ... Es ist nicht so, dass sie im Krankenhaus gelandet wäre“).

Aufgrund der unbewussten Natur dieser tief verwurzelten Heuristiken fügten sich viele Australier*innen in ihre Rolle bei der Aufrechterhaltung der Situation und lehnten ihre Verantwortung eher ab. Dies gründete darauf, dass Einflussnahme als kostspielig angesehen wurde, so dass viele schnell bereit waren, sich nicht einzumischen, wenn es zu respektlosem Verhalten zwischen jungen Männern und Frauen kam. Es ist eine verbreitete Norm, diese Erfahrung als wichtigen Teil der Entwicklung junger Menschen zu betrachten. Menschen, die Einfluss nehmen wollen, müssen mehr Selbstvertrauen und Selbstwirksamkeit spüren, um angemessen reagieren zu können.

Das Ergebnis dieser tiefgehenden Evidenzsammlung war die Entwicklung einer langfristigen nationalen Kampagne zur Verhaltensänderung, die darauf abzielte, die tief verwurzelten Einstellungen und Überzeugungen auf gesellschaftlicher Ebene im Laufe der Zeit in Frage zu stellen und zu verändern. Die Kampagne „Stop it at the Start“ ist eine langfristige Investition, die dem Ausmaß und der Komplexität des Themas gerecht wird – denn die Veränderung tief verwurzelter gesellschaftlicher Einstellungen geschieht nicht über Nacht. Es handelt sich um eine Primärpräventionskampagne. Seit ihrem Start wurde sie unabhängig evaluiert und gilt als die erfolgreichste Kampagne, die jemals von der australischen Regierung gestartet wurde.

Sie hat zu „neuen Diskussionen“ über die Zusammenhänge zwischen Respektlosigkeit und Gewalt gegen Frauen geführt, zu einer deutlichen Verringerung der unbewussten Heuristiken, die das Problem unbeabsichtigt aufrechterhalten, und zu einer deutlichen Verringerung der gewaltunterstützenden Einstellungen unter jungen Menschen. Und schließlich ist die durchschnittliche Zahl der australischen Frauen, die von einem aktuellen oder ehemaligen Partner getötet werden, deutlich gesunken, von ein oder zwei pro Woche (2015) auf eine alle 11 Tage (2022).

Welche Art von Evidenz half, die Ursache des Problems zu erkennen?

Qualitative Forschung, um den Kern des Problems zu dekonstruieren: um zu verstehen, warum das Problem in der Gesellschaft fortbestand, obwohl die meisten Australier*innen Gewalt gegen Frauen und Kinder verurteilten. Innovative qualitative Techniken wurden in Kombination mit dem eigenen Behaviour Change Framework von Verian eingesetzt.

Verian sammelte die Perspektiven vieler verschiedener Gruppen in unterschiedlichen Settings: Es wurden Fokusgruppen mit Müttern, Vätern, Eltern zusammen, älteren und jüngeren Geschwistern, Großeltern und nicht-elterlichen Bezugspersonen durchgeführt. Es fanden Tiefeninterviews mit vulnerableren Zielgruppen wie Menschen mit Behinderungen, Mitgliedern der LGBTQIA-Community und Überlebenden häuslicher Gewalt statt.

Themenunabhängige Empfehlungen für die Diagnosephase:

1. Überprüfen Sie zunächst die verfügbaren Daten und identifizieren Sie wichtige Evidenzlücken. Lücken können fehlende Perspektiven der Beteiligten oder fehlende Beschreibungen der Ursachen und Einflüsse des Problems sein.
2. Wenden Sie holistisches Denken an, insbesondere wenn es um ein langfristiges Problem geht, für das es keine offensichtlichen Ursachen und Lösungen gibt. Integrieren Sie Erkenntnisse und Methoden aus den Verhaltenswissenschaften, um zu verstehen, warum reales Verhalten oft vom beabsichtigten Verhalten abweicht.

DESIGN: Erarbeiten und Auswählen von Lösungen

Sind das Problem und seine Ursachen identifiziert, geht es im nächsten Schritt darum, die beste Lösung für dieses Problem zu entwickeln und auszuwählen.

Abhängig von der Art des Problems und dem Grad der Neuheit des Problems wird möglicherweise eine Sondierungsphase für neue Lösungen oder eine Literaturrecherche benötigt, um Best-Practice-Lösungen zu identifizieren. In den meisten Fällen gibt es eine Vielzahl möglicher Lösungen, für die in der Regel nachgewiesen werden muss, welche davon die größte Wirkung und/oder das beste Kosten-Nutzen-Verhältnis bietet. Die folgenden Fragen sind typisch für die Entscheidungsfindung in dieser Phase des Politikzyklus:

DESIGN & ENTSCHEIDUNG Wie kann die beste Lösung für das Problem gefunden werden?	Wie kann man Lösungen finden, um das Problem anzugehen? <ul style="list-style-type: none">> Können wir etwas anpassen, das anderswo funktioniert hat?> Wie entwickelt man neue Lösungen?
	Wie erkennt man effektive oder kostengünstige Lösungen? <ul style="list-style-type: none">> Wie testet und bewertet man die Optionen?> Erreicht eine Maßnahme mehr als eine andere mit der gleichen Investition?> Wie wendet man eine gute „Theory of Change“ oder ein Logik-Modell an?
	Wie kann die Zustimmung der Stakeholder sichergestellt werden? <ul style="list-style-type: none">> Wie kann man Stakeholder angemessen einbeziehen oder deliberative Methoden einsetzen und zusammenarbeiten/mitgestalten?

Das erste Beispiel veranschaulicht die Herausforderungen in der Designphase bei der Suche nach möglichen Lösungen. Es basiert auf den Erfahrungen des schwedischen Nationalen Rates für Kriminalprävention und der Universität Malmö in Schweden, die die Konzeptionsphase mit einer 2020 veröffentlichten Prozessbewertung begleitet haben.



Wie kann man Lösungen finden, um das Problem anzugehen? Wie adaptiert man etwas, das anderswo funktioniert hat?

Schwere Gewaltverbrechen sind eines der größten gesellschaftlichen Probleme in Schweden. Die Hauptquelle dieser Gewalt scheint innerhalb und zwischen kriminellen Gruppen zu liegen und steht im Zusammenhang mit Schusswaffen. Die schwedische Forschung darüber, was zur Verringerung schwerer Gewalt beiträgt, ist begrenzt, und es sind mehr Kenntnisse erforderlich, um der Polizei und anderen Akteuren der Kriminalprävention die Instrumente an die Hand zu geben, die sie benötigen, um den Trend umzukehren.

Um mögliche Strategien zur Bekämpfung tödlicher Gewalt zu identifizieren, untersuchte der schwedische Nationale Rat für Kriminalprävention (Brå) Regionen auf der ganzen Welt, die sich zuvor mit tödlicher Gewalt befasst hatten. Sie fanden heraus, dass eine Strategie namens „Group Violence Interventions“ (GVI) in den 1990er Jahren in Boston und anderen amerikanischen Städten erfolgreich angewendet wurde.

Um mehr über die Anforderungen und die praktische Umsetzung dieser GVI-Strategie zu erfahren, besuchten Vertreter*innen von Brå die Vereinigten Staaten und berieten sich mit Professor David M. Kennedy, der in den 1990er Jahren das Boston Gun Project leitete und Direktor des National Network for Safe Communities in den Vereinigten Staaten geworden war.

Die GVI-Politik, die in Boston so gut funktionierte, wurde von einer interdisziplinären Arbeitsgruppe in einem Schritt-für-Schritt-Prozess entwickelt, der half, mögliche Probleme zu erkennen und effektiv zu lösen (Problemlösungsmodell-Ansatz). Die Arbeitsgruppe bestand aus Praktiker*innen aus verschiedenen Trägern und Behörden sowie aus Akademiker*innen. Gemeinsam kombinierten sie qualitative und quantitative Forschung mit den Erfahrungen von Polizist*innen und Streetworkern.

Der GVI-Ansatz versuchte, Bandengewalt zu verhindern, indem er die Bandenmitglieder davon überzeugte, dass Gewalt und Waffengebrauch Konsequenzen haben würden, was sie dazu veranlasste, ihr Verhalten zu ändern. Ein Schlüsselement der Strategie bestand darin, einer relativ kleinen Zielgruppe eine direkte und explizite "Abschreckungsbotschaft" zu vermitteln, die die Reaktion auf bestimmte Verhaltensweisen offenlegte. Das Boston Police Department hat die Beteiligung an Bandengewalt für die Beteiligten kostspieliger gemacht. Es wurde erwartet, dass die Kommunikation dessen, was mit anderen in der Zielgruppe passiert war, weitere Bandengewalt in der Stadt verhindern würde.

Nach der erfolgreichen Implementierung von GVI in Boston wurde die Strategie vor allem in amerikanischen Städten, aber auch in einigen europäischen Städten wie Glasgow und London umgesetzt. Die Bedingungen für die Einführung von GVI in Schweden wurden mit dem Hinweis auf die Tatsache in Frage gestellt, dass die Unterschiede zwischen den Vereinigten Staaten und Schweden zu groß seien.

Daher wurde beschlossen, ein Pilotprojekt in einer schwedischen Stadt durchzuführen. Dieser Praxisversuch wurde von einer Prozess- und Wirkungs-Evaluation begleitet, um die Elemente zu verstehen, die erforderlich sind, um eine solche Maßnahme an verschiedene Kontexte/Regionen anzupassen. Die Gemeinde Malmö wurde für zwei Jahre zu einem Pilotprojekt für GVI in Schweden. Das amerikanische Team wurde aufgrund seines Wissens konsultiert, das auf seiner Rolle bei der früheren GVI-Arbeit von Boston basiert.

Die 2021 veröffentlichte Wirkungsevaluation dieses Versuchs zeigte, dass die Zahl der Schießereien in Malmö während des Zeitraums des Pilotprojekts zurückging. Es ist jedoch schwierig, statistisch zu ermitteln, welche Faktoren hinter der Reduktion stehen. Nichtsdestotrotz waren die Ergebnisse vielversprechend genug, so dass immer mehr schwedische Städte diese Politik in ihren Gebieten übernahmen und anpassten.

Themenunabhängige Empfehlungen für die Designphase:

1. Prüfen Sie, was anderswo funktioniert hat, aber seien Sie sich bewusst, dass möglicherweise Anpassungen vorgenommen werden müssen.
2. Stellen Sie sicher, dass Sie das zugrunde liegende Wirkungsmodell der Politikmaßnahme verstehen, das unabhängig von den lokalen Anforderungen verwendet werden kann.
3. Stellen Sie ein Team aus Wissenschaftler*innen und Praktiker*innen zusammen, um politische Lösungen zu entwickeln und anzupassen, um die Wahrscheinlichkeit zu erhöhen, dass sie in der Realität funktionieren.

Da es nicht immer möglich oder wirtschaftlich sinnvoll ist, Experimente in der realen Welt durchzuführen, um die Wirksamkeit einer oder mehrerer Lösungen zu testen, veranschaulicht das nächste Beispiel, wie ein solcher Test schneller und kostengünstiger wie in einem Laborexperiment durchgeführt werden kann. Daher können auch mehrere Lösungen oder Variationen von Lösungen verglichen werden. Das Beispiel basiert auf den Erfahrungen bei Verian im Vereinigten Königreich.



Wie erkennt man effektive Lösungen? Wie lassen sich politische Maßnahmen in der Designphase simulieren und testen?

Laut Ofcom, der britischen Regulierungsbehörde für Kommunikation, glauben die meisten Internetnutzenden, dass die Vorteile des Internets die Risiken überwiegen. Sieben von zehn Nutzenden geben jedoch an, in den letzten drei Monaten etwas potenziell Schädliches online gesehen oder erlebt zu haben. Online-Plattformen verwenden häufig Sicherheitsmaßnahmen wie Warnmeldungen, Meldemechanismen, elterliche Kontrollmöglichkeiten und allgemeine Geschäftsbedingungen, um ihre Nutzer zu schützen. Es gab jedoch nur begrenzte öffentliche Forschung über deren Wirksamkeit. Gleichzeitig gibt es immer mehr Erkenntnisse darüber, dass relativ kleine Änderungen am Design der Online-Umgebung, die Sicherheitsfunktionen beinhalten könnten, das Nutzerverhalten beeinflussen können – zum Guten und zum Schlechten.

Um die Auswirkungen bestimmter Warnmeldungen und Mechanismen zur Meldung von Inhalten auf Video-Sharing-Plattformen (VSPs) zu untersuchen, wurde ein randomisiertes, kontrolliertes Experiment (RCT) in eine Online-Umfrage eingebettet, die eine natürliche VSP (wie YouTube, Vimeo oder TikTok) simulierte. Dieses Laborexperiment ermöglichte die Beobachtung typischer Verhaltensweisen beim Scrollen durch einen Social-Media-Feed. Jede*r Befragte interagiert mit einer Version einer VSP-Oberfläche und hatte die Möglichkeit, eine Reihe von kurzen Videos anzusehen: einige mit neutralen Inhalten und andere mit legalen, aber potenziell schädlichen Inhalten. Die Befragten wurden nach dem Zufallsprinzip entweder einer Kontrollgruppe oder einer Experimentalgruppe zugeteilt. Der Kontrollgruppe wurde eine Oberfläche ohne Warnmeldung angezeigt. Jede der vier Experimentalgruppen erhielt eine andere Art von Warnmeldungen.

Die Ergebnisse deuten darauf hin, dass Warnmeldungen die Befragten dazu veranlassten, sich stärker mit Entscheidungen darüber zu befassen, inwieweit sie die ihnen präsentierten Videoinhalte anschauen möchten, aber nur die hochgradig deskriptive Social-Proof-Warnmeldung "Dieses Video enthält Material, das andere Zuschauer auf dieser Plattform als sensibel gemeldet haben" hatte einen statistisch signifikanten Einfluss darauf, ob die Befragten legale, aber potenziell schädliche Inhalte übersprungen haben. Diese Form der Warnmeldung gab den Befragten, die den Inhalt nicht sehen wollten, die Informationen, die ihnen die Entscheidung ermöglichten, ihn zu überspringen.

Ebenso waren die Befragten, die sich dafür entschieden, den Inhalt anzusehen, nachdem sie einer Warnmeldung ausgesetzt waren, besser in der Lage, das Schadensrisiko zu verstehen: Sie hatten entweder mit dem Inhalt keine Probleme oder waren sich der Tatsache bewusst, dass sie den Inhalt problematisch finden könnten, und waren daher darauf vorbereitet, ihn bei Bedarf zu überspringen. Die effektivste Lösung in diesem Online-Experiment zeigte auch keine negativen Auswirkungen, da sie die Wahrscheinlichkeit, neutrale Inhalte zu überspringen, nicht erhöhte und nicht als störend empfunden wurde.

Für Ofcom war dies die erste Erfahrung mit einem Online-RCT, bei dem eine nachgebildete Benutzeroberfläche verwendet wurde, um typische (und oft unbewusste) Entscheidungsszenarien zu simulieren. Die Studie hat gezeigt, dass eine gute Simulation auch dann ein gutes Gleichgewicht zwischen hochwertiger Evidenz in kurzer Zeit und einem akzeptablen Budget bieten kann, selbst wenn die Tests nicht in einer realen Umgebung stattfinden. Diese experimentellen Ergebnisse sind realistischer als eine einfache Befragung mit hypothetischen Fragen zum Verhalten. Eine größere Stichprobengröße ermöglicht es, mehrere Varianten zu testen und zu vergleichen. Das experimentelle Labor-Design ermöglicht es außerdem, alle anderen externen Einflüsse zu kontrollieren, um die Auswirkungen der verschiedenen Bedingungen zu isolieren (was in einem real-world Design nicht möglich ist). Ofcom möchte daher diesen methodischen Ansatz weiterführen.

Themenunabhängige Empfehlungen für die Designphase:

1. Wenn Sie verschiedene Konzepte/Versionen testen, sollten Sie RCTs (randomisierte, kontrollierte Experimente) in Betracht ziehen, die die Entscheidungssituation mit realistischem Kontext und realitätsnaher Entscheidungsstruktur simulieren, um realistisches Verhalten zu beobachten.
2. Stellen Sie sicher, dass sie alle potenziell beeinflussenden Faktoren in der Studie berücksichtigen, um sie zu kontrollieren und die Wirkung der Konzepte besser zu verstehen.
3. Um langfristige Auswirkungen der Konzepte auf das Verhalten zu messen, ist es notwendig, Follow-up-Befragungen zur Beobachtung von wiederholtem Verhalten oder anhaltenden Einstellungsänderungen zu implementieren.

Das nächste Beispiel veranschaulicht die Herausforderungen, um die Zustimmung der Stakeholder in der Designphase sicherzustellen, wenn eine Lösung ausgewählt werden muss, die stark umstritten ist. Es basiert auf den Erfahrungen von Verian in Frankreich.



Wie kann die Zustimmung aller Stakeholder sichergestellt werden? Was ist zu tun, wenn starker politischer Druck auf bestimmte politische Optionen ausgeübt wird?

Die französische Regierung hat sich das ehrgeizige Ziel gesetzt, die Recyclingquote von Kunststoffflaschen von 61 % im Jahr 2021 auf 90 % bis 2029 zu erhöhen. Darüber hinaus strebt das Land an, den Verbrauch von Plastikflaschen bis 2030 zu halbieren. Um diese neuen Ziele zu erreichen, sah der Minister für den ökologischen Wandel die Einführung eines neuen Pfandsystems für das Recycling von Plastikflaschen vor.

Dieses Vorhaben war jedoch höchst umstritten. Nach Ansicht seiner Befürworter, der Getränkeindustrie und Umweltorganisationen, ist ein Pfandsystem ein wichtiger Hebel, um die von der EU gesetzten Recyclingziele zu erreichen. Andere Länder, wie z.B. Deutschland, zeigen, dass ein solches System bei der Erreichung der Recyclingziele wirksam sein kann.

Gegner des Vorhabens waren lokale Behörden, die bereits stark in das derzeitige Recyclingsystem investiert hatten. Eine Umstellung auf ein Pfandsystem würde ihnen eine Einnahmequelle entziehen.

Es wurde auch befürchtet, dass die wertvollsten Abfälle zur Wiederverwertung an die Hersteller zurückgegeben würden, die lokalen Behörden jedoch auf den wertlosen Abfällen sitzen blieben, die am schwierigsten zu recyceln sind. Es war daher ein glücklicher Zufall, dass ein 2020 verabschiedetes Gesetz den Minister verpflichtete, seine politische Entscheidung auf solide Evidenz zu stützen.

Die französische Agentur für den ökologischen Wandel (ADEME) beauftragte daraufhin ein Forschungskonsortium (Verian, Ifop und EY), Erkenntnisse zu sammeln, um das Konzept eines neuen Pfandsystems für das Recycling von Getränkeflaschen im Vergleich zum bestehenden Recyclingprozess zu bewerten. Die Evaluation stützte sich auf eine Reihe von Datenquellen: europäisches Benchmarking von Pfandsystemen, eine quantitative Verbraucherbefragung zum Kauf- und Sortierverhalten, RCTs zur Prüfung der Wahrnehmungen und Praktiken der Verbraucher für das neue Konzept im Vergleich zum Status quo und eine wirtschaftliche Analyse zur Untersuchung der Wirtschaftlichkeit des neuen Konzepts im Vergleich zum aktuellen.

Die im Juni 2023 veröffentlichten empirischen Ergebnisse zeigten, dass mit dem neuen Pfandsystem nicht alle gesetzten Ziele vollständig erreicht werden würden. Darüber hinaus wäre das derzeitige System auch in der Lage, die Ziele zu erreichen, vorausgesetzt, dass alle geplanten Ressourcen eingesetzt würden.

Angesichts dieser Schlussfolgerungen, aber auch (vielleicht vor allem) angesichts des starken Widerstands der lokalen Behörden gegen das neue Pfandsystem, beschloss der Minister für den ökologischen Wandel im September 2023, den Plan eines Pfandsystems für Kunststoffverpackungen aufzugeben. Er kündigte jedoch an, dass dies durch die Festlegung ambitionierter Ziele für die lokalen Behörden zur Verbesserung der Sammlung und des Recyclings ausgeglichen werden solle, insbesondere auf der Grundlage der von der ADEME identifizierten Strategien.

Dank dieser evidenzbasierten Entscheidungsfindung waren die Reaktionen der Stakeholder und der Medien überwältigend positiv: "Die Regierung weicht von einer gut-gemeinten, aber falschen Idee ab", "ein Sieg für unsere Regionen, für unsere Bürgerinnen und Bürger und im Kampf gegen den Klimanotstand angesichts der budgetären und ökologischen Auswirkungen, die diese Maßnahme mit sich gebracht hätte".

Themenunabhängige Empfehlungen für die Designphase:

1. Themen, die unter hohem politischen Druck stehen, benötigen besonders vertrauenswürdige und unabhängige Quellen und Evidenz.
2. Stellen Sie sicher, dass alle relevanten Stakeholder gehört und die wichtigsten Argumente geprüft werden.
3. Evidenz kann Entscheidungsträgern helfen, Akzeptanz für ihre Entscheidungen zu schaffen.

IMPLEMENT: Feinabstimmung der Implementierung einer Lösung

Ist die Entscheidung gefallen, welche Lösung für ein Problem implementiert werden soll, geht es in der nächsten Phase um die Feinabstimmung während der Implementierungsphase, einschließlich der Planung der Kommunikation.

Diese Phase wird von politischen Entscheidungsträger*innen oft vernachlässigt, da sie sich in der Regel auf die Makroebene einer politischen Lösung konzentrieren. Der Erfolg und die Wirksamkeit eines Vorhabens hängen jedoch oft von einer gut konzipierten Umsetzung und Kommunikation in der Praxis ab. Um sicherzustellen, dass Politik so effizient und effektiv wie möglich durchgesetzt wird, ist es wichtig, dass das Umsetzungskonzept sorgfältig hinterfragt und getestet wird, idealerweise mit allen relevanten Intermediären und Zielgruppen. Typische Forschungsfragen in dieser Phase sind:

IMPLEMENTIEREN & KOMMUNIZIEREN

Sollten wir zusätzliche Schritte unternehmen, um die Wahrscheinlichkeit zu erhöhen, dass die ausgewählte Lösung das tut, was wir beabsichtigen?

Was könnte dazu führen, dass die Maßnahme bei der Umsetzung scheitert? Wie kann ein Vorhaben mit zusätzlichen Erkenntnissen für eine Kommunikationsstrategie verfeinert werden?

Wie kann die Partizipation von Intermediären und Zielgruppen über einen längeren Umsetzungszeitraum sichergestellt werden?

Wie skaliert man Interventionen, die sich während eines Pilotprojekts als wirksam erwiesen haben?

Das erste Beispiel zeigt, wie Widerstände gegen Politikmaßnahmen und Programmen durch Kommunikationsstrategien überwunden werden können. Es basiert auf den Erfahrungen von Verian in den Niederlanden.



Was könnte dazu führen, dass die Maßnahme bei der Umsetzung scheitert? Wie kann ein Vorhaben mit zusätzlichen Erkenntnissen für eine Kommunikationsstrategie verfeinert werden?

Die Niederlande sind als "Land der Fahrräder" bekannt und haben die höchste Rate an Fahrradbesitzenden weltweit. Gleichzeitig verzeichnete das Land einen bemerkenswerten Anstieg verunglückter Radfahrer*innen. Ein Schlüsselfaktor für die Verbesserung der Sicherheit ist die Verwendung von Fahrradhelmen, aber nur sehr wenige Menschen in den Niederlanden tragen einen Helm bei regelmäßigen Fahrradtouren.

Als Reaktion darauf wollte das Ministerium für Infrastruktur und Wasserwirtschaft einen umfassenden Fahrradsicherheitsplan entwickeln, um die Zahl der schweren Verletzungen unter Radfahrer*innen zu reduzieren. Um jedoch von Bürger*innen akzeptiert zu werden, die überwiegend gegen das Tragen von Helmen zu sein schienen, war eine Studie darüber erforderlich, was die Einstellungen, Überzeugungen und Verhaltensweisen der niederländischen Bürger*innen in Bezug auf das freiwillige Tragen von Helmen bei regelmäßigen Fahrradtouren beeinflusst.

Nur mit dem Wissen um die zentralen Antriebsfaktoren des Widerstands gegen das Tragen von Helmen können wirksame Kommunikations- und Umsetzungsmaßnahmen konzipiert und politische Maßnahmen zur Helmnutzung von den Bürgerinnen und Bürgern auch tatsächlich verfolgt werden.

Da Einflussfaktoren bei einem Thema, bei dem die vorherrschende soziale Norm so stark ausgeprägt ist, schwer zu bestimmen sind, konzipierte Verian ein zweistufiges Forschungsprojekt. Interviews mit Eltern von Kleinkindern, Pendlern und Senior*innen, die regelmäßig Fahrrad fahren, ergaben, dass die Haupthindernisse für die Nutzung von Helmen eine geringe Risikowahrnehmung und die Überzeugung waren, dass Helme unnötig sind und sich negativ auf das Image auswirken. Zu den wichtigsten Hindernissen gehörten soziale Normen, Stereotypen, Probleme mit dem Aussehen und Unbequemlichkeit. Die Studie zeigte jedoch auch, dass die Akzeptanz gesteigert werden könnte, wenn der Helm häufiger getragen wird, die Risikowahrnehmung steigt, sich Design und Komfort verbessern und persönliche Werte wie Verantwortung und Gesundheitsbewusstsein stärker mit der Helmnutzung in Verbindung gebracht werden.

Die quantitative Befragung ergab, dass ein beträchtlicher Teil der Nicht-Helmträger*innen offen dafür wäre, in Zukunft einen Helm zu tragen. Um diese und andere Radfahrer*innen anzusprechen, entwickelte das Ministerium eine Kommunikationsstrategie zur Förderung des freiwilligen Helmtragens unter Radfahrer*innen, die sich auf die eher impliziten, emotionalen Faktoren konzentriert, die nicht nur rationale Kosten-Nutzen-Berechnungen berücksichtigen. Die Studie veranlasste das Ministerium, eine ganzheitliche Sicht auf die Entscheidungsfindung der Bürger*innen in Bezug auf die Verwendung eines Helms einzunehmen und sicherzustellen, dass rationale Entscheidungen nicht durch emotionale Widerstände dominiert werden.

Themenunabhängige Empfehlungen für die Implementierungsphase:

1. Nehmen Sie eine holistische Sicht auf das Umfeld ein, in der eine Maßnahme umgesetzt werden soll, und berücksichtigen Sie sowohl emotionale als auch rationale Reaktionen.
2. Unterschätzen Sie nicht die Stärke von Gewohnheiten und sozialen Normen, wenn eine Politikmaßnahme Verhaltensänderungen erfordert.
3. Entwickeln Sie eine durchdachte Kommunikationsstrategie, die sich bei der Einführung neuer Initiativen nicht nur auf rationale Überlegungen stützt, sondern auf Erkenntnissen darüber basiert, wie und auf was die Bürger*innen reagieren werden.

Das nächste Beispiel verdeutlicht die Herausforderung, die Akzeptanz der Akteure in der Umsetzungsphase sicherzustellen. Es basiert auf den Erfahrungen von Ramboll in Deutschland.



Wie kann die Partizipation von Intermediären und Zielgruppen über einen längeren Umsetzungszeitraum sichergestellt werden?

Um die Lehrerbildung im Studium zu verbessern, hat das Bundesministerium für Bildung und Forschung (BMBF) das große Förderprogramm "Qualitätsoffensive Lehrerbildung" ins Leben gerufen. Ziel des Programms war es, die Bildungsstrukturen an den Hochschulen zu optimieren, relevante Unterstützungsangebote für Studierende zu entwickeln und die Vergleichbarkeit von Studium und Qualifikationen über die Bundesländer hinweg sicherzustellen.

Das Bundesministerium beauftragte Ramboll Management Consulting und die Universität Linz mit einer umfassenden Evaluation, um die Wirkung dieses groß angelegten Programms zu bewerten.

An dem Programm waren 72 Hochschulen aus allen Bundesländern mit 91 sehr heterogenen Projekten beteiligt. Daher musste eine große und vielfältige Gruppe von Akteuren und Interessengruppen einbezogen und über einen relativ langen Evaluierungszeitraum zur Teilnahme motiviert werden.

Die Evaluatoren stellten die Zusammenarbeit der Akteure in den Mittelpunkt, und das Programm zielte auf eine länderübergreifende Zusammenarbeit ab, um Wirkung und Vergleichbarkeit zu gewährleisten.

Um die Bereitschaft aller Stakeholder zur Beteiligung an der Datenerhebung über den gesamten Zeitraum (2016 – 2024) aufrecht zu erhalten, hat Ramboll ein groß angelegtes Evaluationsprogramm aufgesetzt, das Online-Befragungen für alle Projekte, Monitoringsysteme, Experteninterviews, Interviews mit geförderten und nicht geförderten Hochschulen sowie Gruppendiskussionen mit Vertreterinnen und Vertretern der Bildungs- und Wissenschaftsministerien der Bundesländer umfasste. Außerdem war die Beobachtung von Veranstaltungen und eine Dokumentenanalyse vorgesehen. Diese große Vielfalt an Datenquellen stellte sicher, dass alle relevanten Daten erhoben werden konnten, ohne die einzelnen Teilnehmenden zu stark zu belasten.

Außerdem konnten sie jeweils mit einer geeigneten Forschungsmethode angesprochen werden. Sehr heterogene Ausgangspositionen an den Hochschulen wurden durch umfangreichere und weniger standardisierte Monitoringinstrumente (z. B. mehr offene Antwortmöglichkeiten) berücksichtigt. Die Unterstützung von Vertreter*innen der Bundesländer und eines wissenschaftlichen Beirats untermauerte die Expertise und Autorität der Forschenden. Darüber hinaus spielte die Transparenz der Evaluationsergebnisse, z. B. durch jährliche thematische Zwischenberichte und Workshops zur Erörterung von Zwischenergebnissen, eine Schlüsselrolle bei der Einbindung von Stakeholdern, indem sie das Vertrauen stärkte und den Universitäten bereits einige Vorteile bot. Diese Maßnahmen sorgten dafür, dass die „Qualitätsoffensive Lehrerbildung“ Reformen unterstützte, neue Wege in der Lehrerbildung eröffnete und die Forschung in der Lehrerbildung stärkte. Die Projekte und Stakeholder wurden bundesweit vernetzt und in ihrer Arbeit unterstützt. Die „Qualitätsoffensive Lehrerbildung“ half den Hochschulen, den Abstimmungsbedarf besser zu bewältigen und durch Vernetzung resilienter zu werden.

Themenunabhängige Empfehlungen für die Implementierungsphase:

1. Wenn Sie Akteure und Intermediäre einbeziehen, stellen Sie sicher, dass Ihre Anfrage für sie relevant und leicht zu erfüllen ist.
2. Heben Sie den Nutzen der Begleitforschung für die Stakeholder hervor.
3. Erhöhen Sie das Vertrauen, indem Sie so transparent und offen wie möglich sind.

Das nächste Beispiel veranschaulicht die Herausforderungen bei der Skalierung in der Implementierungsphase nach einem erfolgreichen Pilotprojekt. Die Fallstudie basiert auf den Erfahrungen der Behavioural Insights Unit von New South Wales (NSW) mit der Erprobung einfacher Verhaltensinterventionen, um die geschlechtsspezifische Lücke bei der Rekrutierung für die NSW-Regierung in Australien zu verringern.



Wie skaliert man Interventionen, die sich während eines Pilotprojekts als wirksam erwiesen haben?

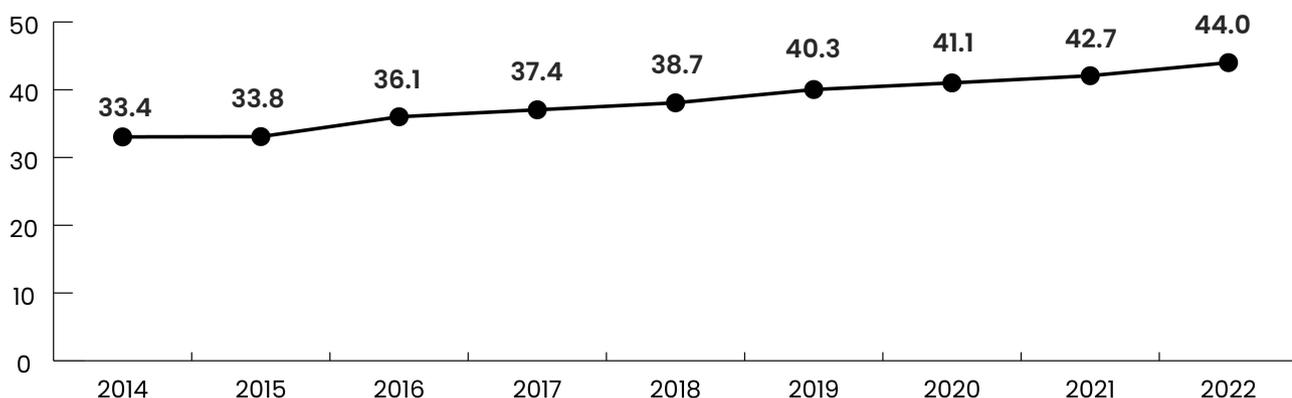
Seit 2015 war es eine der Prioritäten des Premiers von New South Wales, den Anteil von Frauen in leitenden Führungspositionen im NSW-Regierungssektor bis 2025 von 33 % auf 50 % zu erhöhen. Eine zentrale Herausforderung bestand darin, dass sich Frauen nicht im gleichen Maße wie Männer für leitende Führungspositionen bewerben (dies wurde in der Diagnosephase entdeckt). Das Behavioural Insights-Team entwickelte zwei einfache, kostengünstige Interventionen, die Frauen ermutigen sollen, sich erneut für leitende Positionen zu bewerben, nachdem sie knapp gescheitert waren.

Die Interventionen bestanden aus einer E-Mail und einem Telefonat mit einem/einer Personalverantwortlichen, das sich darauf konzentrierte, wie gut der/die Bewerber*in abgeschnitten hatte, und sie ermutigte, sich erneut zu bewerben.

Beide Interventionen nutzen die Verhaltenstendenz, eine Erfahrung danach zu beurteilen, wie wir uns während ihres emotionalen Höhepunkts und am Ende fühlen. Während des Rekrutierungsprozesses ist der emotionale Höhepunkt oft ein Vorstellungsgespräch, während das Ende für die meisten Bewerber*innen eine generische E-Mail ist, in der ihnen mitgeteilt wird, dass ihre Bewerbung nicht erfolgreich war. Die Intervention zielte darauf ab, diesen finalen Eindruck aufzubrechen, indem Bewerber*innen, die ein Vorstellungsgespräch erhielten, aber nicht erfolgreich waren, eine ermutigende E-Mail oder einen Telefonanruf erhielten. Die Interventionen wurden 2019/2020 in einem Pilotprojekt mit einer Gruppe von Stellen im öffentlichen Dienst getestet. Es wurde ein randomisiert-kontrolliertes Studiendesign mit 1.614 Bewerber*innen implementiert. Diese wurden nach dem Zufallsprinzip so zugeteilt, dass sie entweder einen "Business as usual"-Rekrutierungsprozess oder eine ermutigende E-Mail oder eine ermutigende E-Mail plus einen Telefonanruf erhielten. Zwei Wochen nach Erhalt der Intervention (bzw. für die Kontrollgruppe ohne Intervention) erhielten die Bewerber*innen eine Umfrage, um ihre allgemeine Zufriedenheit mit dem Einstellungsverfahren, ihre Absicht, sich erneut zu bewerben, und die Wahrscheinlichkeit, die NSW-Regierung zu empfehlen, zu erfassen. Darüber hinaus wurde geprüft, ob sich die Bewerber*innen innerhalb von drei und sechs Monaten erneut für eine andere leitende Position in der NSW-Regierung bewarben.

Die Pilotergebnisse zeigten deutlich, dass beide Interventionen erfolgreich waren. Sie hatten nicht nur einen erheblichen Einfluss auf die Absicht, sondern auch auf das Verhalten, sich erneut für leitende Positionen zu bewerben. Darüber hinaus war der Einfluss bei Frauen signifikant größer als bei Männern, so dass sich der Gender Gap im Vergleich zur Kontrollgruppe mit einem Gender Gap von 45% signifikant auf 12% (nur E-Mail) bzw. auf 4% (Telefon und E-Mail) reduzierte. Auf der Basis dieser Ergebnisse wurde berechnet, dass, wenn die E-Mail und der Telefonanruf auf die gesamte NSW-Regierung ausgeweitet werden würden, dies den Anteil von Frauen in leitenden Positionen jedes Jahr um bis zu 0,5 % erhöhen und einen wesentlichen Beitrag zur Erreichung einer gleichberechtigten Vertretung der Geschlechter in leitenden Positionen leisten könnte. Aus diesem Grund entwickelte die Kommission für den öffentlichen Dienst in NSW ein Toolkit, um Personalverantwortliche bei der Umsetzung der Interventionen zu unterstützen. Seitdem hat sich die Geschlechtervielfalt in den Regierungsgremien von NSW weiter verbessert, obwohl noch nicht evaluiert wurde, inwieweit dieser Fortschritt auf der Einführung der Intervention beruht.

Prozentsatz von weiblichen Führungskräften im öffentlichen Sektor von NSW.



Quelle: NSW Public Sector Commission 2022, Workforce Profile Report 2022, NSW Government

Themenunabhängige Empfehlung für die Implementierungsphase:

1. Die Skalierung von Interventionen, die sich während eines Pilotprojekts als wirksam erwiesen haben, ist immer schwierig, da sie einen Systemwechsel erfordern. Es ist wichtig, die Zustimmung für das Pilotprojekt von denjenigen einzuholen, die an der Skalierung beteiligt sein werden. Die Kosten für die Durchführung solcher Pilotprojekte sind gering, wenn administrative Daten für die Evaluation verwendet werden, aber Behörden verfügen nicht immer über gute Datensysteme, und für die Durchführung gründlicher Evaluationen ist ein gewisses technisches Know-how erforderlich.

EVALUATE: Bewertung von Leistungen, Wirkungen und langfristigen Effekten einer Lösung

Ohne jegliches Monitoring und Evaluation nach der Implementierung eines Vorhabens weiß man nicht, ob und wie die Lösung funktioniert hat. Das Einholen von Evidenz während des Implementierungsprozesses dient als Frühwarnsystem, wenn etwas schief geht und korrigiert werden muss. Wirkungsevaluationen nach der Implementierung liefern dann die Zahlen, die möglicherweise benötigt werden, um das Vorhaben als Erfolgsgeschichte zu erzählen, um das Kosten-Nutzen-Verhältnis zu überprüfen und um zukünftige politische Entscheidungen in diesem Bereich zu treffen. Typischerweise werden in dieser Phase folgende Forschungsfragen gestellt:

EVALUATE

Was hat funktioniert und was nicht und warum?

Wurden mit der gewählten Option die gewünschten Ziele erreicht?

- > Wurden die anvisierten Zielgruppen ausreichend erreicht?
- > Gab es unbeabsichtigte Effekte?
- > War das Kosten-Nutzen-Verhältnis angemessen?

Wie kann man Monitoring- und Evaluationsdaten nutzen, um Prozesse kontinuierlich zu verbessern und nachhaltige Veränderungen zu erreichen?

Das erste Beispiel veranschaulicht die Herausforderung, in der Evaluationsphase umfassende und hilfreiche Forschungsprogramme zu konzipieren. Es basiert auf den Erfahrungen von Verian in Deutschland.



Wurden mit der gewählten Option die gewünschten Ziele erreicht?

Ziel der Bildungsprämie war es, die Teilnahme an beruflichen Bildungsangeboten zu erhöhen, insbesondere für diejenigen, die durch finanzielle Barrieren daran gehindert wurden. Durch das Angebot finanzieller Unterstützung sollte die Weiterbildung für Menschen mit niedrigem Einkommen erschwinglich werden, indem sowohl die finanziellen Kosten als auch andere Hindernisse wie begrenzte Zeit und eingeschränkte Fähigkeiten reduziert werden.

Um zu beurteilen, ob dies erreicht wurde, führte ein Forschungskonsortium bestehend aus Verian, dem Institut für angewandte Wirtschaftsforschung e.V. und dem Forschungsinstitut Betriebliche Bildung (f-bb) eine groß angelegte Evaluationsstudie mit einer ganzheitlichen Betrachtung von Forschungsaspekten und -fragen durch.

Während sich die Evaluation auf die Standardfrage zu Wirksamkeit und Zielerreichung – “Hat das Programm die Teilnahme an beruflichen Bildungsangeboten erhöht?” – konzentrierte, wurde das Forschungsfeld darüber hinaus umfassender betrachtet und auch berücksichtigt:

- Analyse und Beurteilung der Wirtschaftlichkeit
- Analyse und Bewertung des Bildungserfolgs
- Identifikation von Verbesserungspotenzialen bzw. Fehlersteuerung im Prozess
- Weiterentwicklung von Maßnahmen und Steuerungsmaßnahmen für das Bundesprogramm

Ziel der Studie war es auch, die langfristigen Auswirkungen und Erkenntnisse für zukünftige Bildungsförderprogramme zu evaluieren. Dabei nahm die Studie einen breiteren Blick ein und erfasste nicht nur das Weiterbildungsverhalten und die Einstellungen der Teilnehmenden, sondern auch die breitere Relevanz von geförderter Weiterbildung und deren Wirkung auf den sozialen Ausgleich, Mobilisierung von benachteiligten Gruppen und den Abbau ungleicher Bildungsbeteiligung.

Sie befasste sich daher mit der Einbettung des Programms in den gesamten gesellschaftlichen Kontext und mit weiteren relevanten Entwicklungen. Diese breite Fokussierung trug dazu bei, eine Reihe von Verbesserungsmöglichkeiten für eine nachhaltige bedarfsorientierte Förderung der beruflichen Entwicklung zu identifizieren, zeigte aber auch Grenzen der Wirksamkeit des Programms auf. Es zeigte sich, dass die Teilnehmenden mit größerer Wahrscheinlichkeit an qualitativ hochwertigeren Schulungen oder zusätzlichen Schulungen teilnahmen, und insbesondere Personen mit geringerem Einkommen schienen ermutigt zu werden. Außerdem berichteten die Teilnehmenden von einer hohen Zufriedenheit mit der Bildungsprämie. Weitere Analysen zeigten jedoch, dass mögliche Effekte der Bildungsprämie auf einen Selektionsbias zurückgeführt werden könnten, da nur motiviertere Teilnehmende überhaupt teilnahmen. Auch die Auswirkungen auf den wirtschaftlichen und beruflichen Aufstieg waren bescheiden. Unter dem Gesichtspunkt der Wirtschaftlichkeit stellte sich heraus, dass das Programm die Einzelförderung sehr teuer machte.

Eine gut konzipierte Evaluationsstudie mit vielfältigen und umfangreichen Forschungsfragen schaffte es daher, eine ausgewogene und differenzierte Bewertung der Gesamtwirkungen des Programms zu erzielen.

Themenunabhängige Empfehlungen für die Evaluationsphase:

1. Achten Sie darauf, dass Sie alle relevanten Aspekte eines Programms und seiner potenziellen Auswirkungen erfassen, wenn Sie beurteilen, ob es wirksam war.
2. Bei der Bewertung der Langzeitwirkungen sollten auch der soziale Kontext und das breitere Umfeld berücksichtigt werden.

Das nächste Beispiel zeigt, wie Monitoring-Systeme in der Evaluationsphase nützlich sein können, um Trends und neue Entwicklungen im Blick zu behalten. Es basiert auf den Erfahrungen von Verian in den Niederlanden.



Wie kann man Monitoring- und Evaluationsdaten nutzen, um kontinuierlich zu verbessern und nachhaltige Veränderungen zu erreichen?

Der niederländische Nationale Sportverband (NOC*NSF) hat sich zum Ziel gesetzt, die Sportbeteiligung in den Niederlanden deutlich zu erhöhen. Um sicherzustellen, dass alle Programme auf Evidenz beruhen und neue Strategien sofort evaluiert werden, beauftragte der Verband Verian mit der Durchführung einer monatlichen Monitoring-Umfrage, die sich an eine repräsentative Stichprobe der niederländischen Bevölkerung richtete.

Die Umfrage misst jeden Monat die sportliche Beteiligung und das Ausmaß der körperlichen Aktivität sowie die am häufigsten ausgeübten Sportarten und die demografische Zusammensetzung der niederländischen Sportler*innen. Das Monitoring stützt sich auf einen standardisierten Fragenkatalog, der auf einer validierten Methodik basiert, um die Teilnahme am Sport und die körperliche Aktivität zu bewerten. Sie kann somit einen Schlüsselindikator für die Sportbeteiligung liefern, der über die Zeit und über verschiedene Kontexte hinweg verglichen werden kann. Da die Daten monatlich erhoben werden, ergibt sich ein detailliertes Bild von Trends und Mustern, woraus sich Handlungsempfehlungen für die Anpassung von Politikmaßnahmen ableiten lassen. Darüber hinaus ermöglichen die großen Stichproben, die in diesem Zeitraum zusammengetragen wurden, detaillierte statistische Analysen und soziodemografische Vergleiche.

Die Monitoring-Umfrage sieht auch Ad-hoc-Fragen und spezielle Schwerpunkte vor, um gezielt Daten über sich abzeichnende Trends, saisonale Aktivitäten oder andere besondere Interessengebiete zu erheben. Der Monitor kombiniert daher ein standardisiertes Basismodul, das zuverlässige und konsistente Ergebnisse liefert, mit der Möglichkeit, Strategien flexibel anzupassen und auf sich abzeichnende Entwicklungen im Sportbereich zu reagieren.

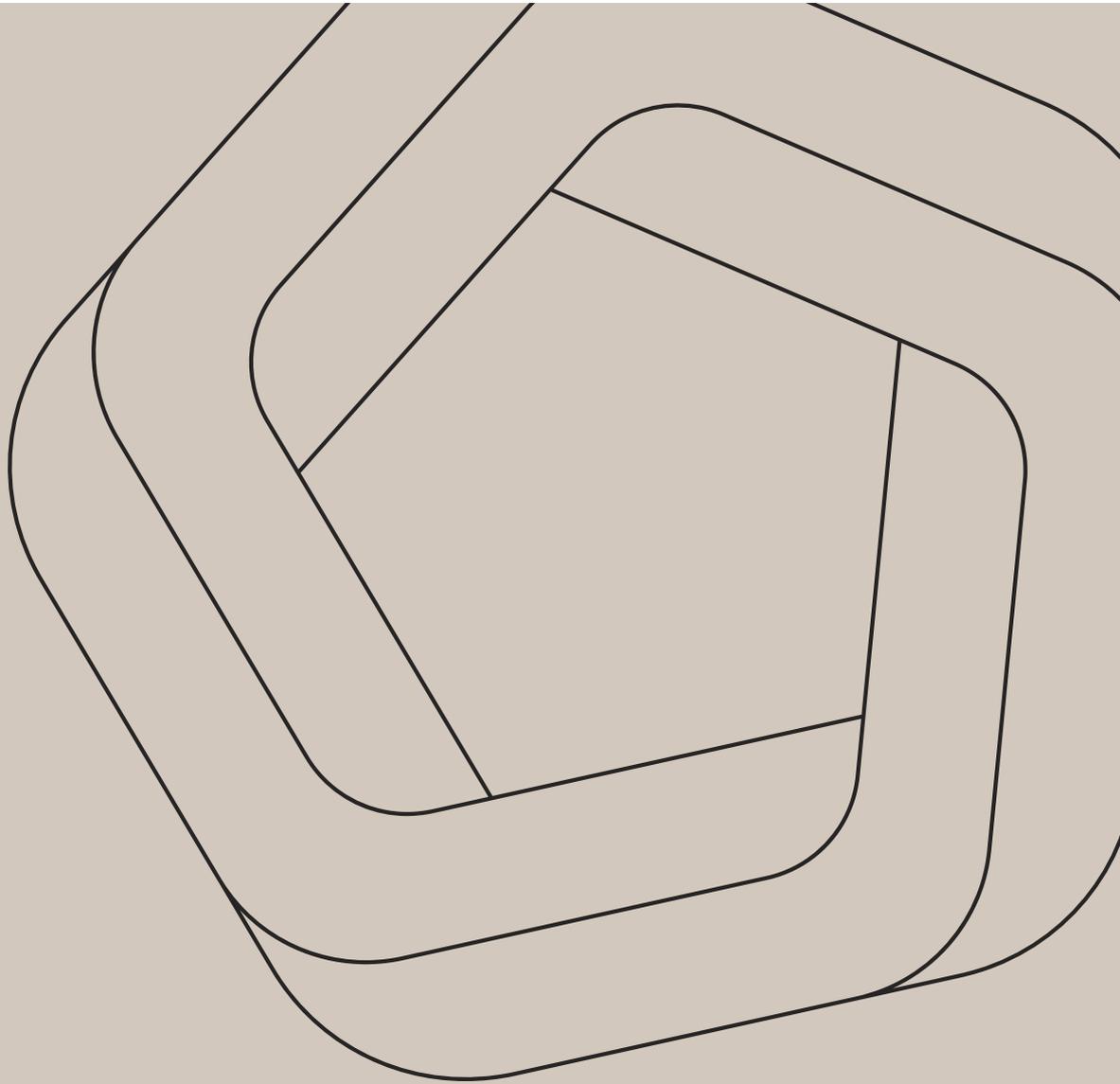
Eine feste Datenpipeline speist die Daten nur wenige Tage nach Beendigung der Feldarbeit in leicht zugängliche Dashboards ein und stellt alle Erkenntnisse sofort zur Verfügung. Das Dashboard ist sehr flexibel und berücksichtigt Feedback aus den Daten der Vormonate, um sicherzustellen, dass es immer auf dem neuesten Stand ist und zeitnahe und korrekte Ergebnisse liefert. Im Allgemeinen ist der Monitor die Grundlage für politische Positionspapiere von NOC*NSF und für die Erstellung langfristiger Strategien, die auf die in der Umfrage identifizierten Barrieren für eine Vielzahl von Untergruppen niederländischer Bürger*innen abzielen.

Themenunabhängige Empfehlungen für die Evaluationsphase:

1. Berücksichtigen Sie bei der Einrichtung von Monitoringsystemen sowohl standardisierte und vergleichbare Fragestellungen als auch flexible Module, um auf Trends und neue Schwerpunkte reagieren zu können.
2. Ein Design mit mehreren Erhebungszeitpunkten hilft beim Verständnis von dynamischen Entwicklungen und liefert große Stichproben für eine detailliertere Analyse.

Literaturhinweise

- Agence de la Transition Écologique (ADEME). (2023). Scénarios avec et sans consigne pour recyclage des emballages de boisson (Scenarios with and without deposit for beverage packaging recycling). Evaluation report, published by ADEME.
- Hadorn, S., Sager, F., Mavrot, C. et al. (2022). Evidence-Based Policymaking in Times of Acute Crisis: Comparing the Use of Scientific Knowledge in Germany, Switzerland, and Italy. *Polit Vierteljahresschr*, 63, 359–382.
- Houses of Parliament. (2017). *Parliamentary Research handbook*.
- Ivert, A.K., Mellgren, C., Nilsson, J. (2020). *Processutvärdering av Sluta Skjut (Process evaluation of Cease Fire)*. Evaluation report published by Malmö University.
- Kennedy, D. M. (2001). *Reducing Gun Violence – The Boston Gun Project’s Operation Ceasefire*. Research Report published by U.S. Department of Justice.
- Morestin, F. (2023). *Consigne des bouteilles en plastique: le gouvernement recule devant une “fausse bonne idée” (Plastic bottle deposit: the government backs away from a false good idea)*. Article published by Novethic on 30th September 2023.
- Nationaler Normenkontrollrat – NKR. (2019): *Erst der Inhalt, dann die Paragraphen. Gesetze wirksam und praxistauglich gestalten*.
- Nesta. (2022). *Engaging with evidence toolkit – A practical resource to strengthen capabilities for evidence use and expert engagement*.
- OECD. (2021). *Regulatory Policy Outlook 2021*.
- Ofcom. (2023). *Behavioural insights for online safety: understanding the impact of video sharing platform (VSP) design on user behaviour*. Economic discussion paper series, issue number 4.
- O’Donoghue, K., Moore, A. (2022). *Reducing Sexual Violence – research informing the development of a national campaign*. Commissioned by the Australian Government, Department of Social Services.
- Riphahn, R.T., Schnitzer, M. (2022). *Die Bedeutung wissenschaftlicher Evidenz für die Arbeit des Bundestags*. In *Vierteljahrshefte*, 91(4), 271–272.
- Sager, F., Hadorn, S., Balthasar, A., Mavrot, C. (2021). *Politikevaluation – Eine Einführung*. Springer VS.
- Seidel, K., Verbeek, H., Fessel, S., Meer, F. (2021). *Nutzen von wissenschaftlicher Evidenz – Erwartungen an wissenschaftliche Expertise*. Bericht über eine Bundestagsbefragung im Rahmen der Leopoldina-Evidenzinitiative unter Leitung von Regina R. Riphahn und Monika Schnitzer. *Diskussion Nr. 27, Halle (Saale)*.
- Social Policy Evaluation and Research Unit (Superu). (2018). *Making sense of evidence: A guide to using evidence in policy*. Wellington, New Zealand.
- World Health Organisation. (2021). *Evidence, policy, impact: WHO guide for evidence-informed decision-making*.



2025 Verian Group. All information and data contained in this document is the property of Verian Group, and cannot be reproduced without our prior permission.