

211-043

DGUV Information 211-043



Gute Praxis der Evaluation von Präventionsmaßnahmen in der gesetzlichen Unfallversicherung

kommmit**mensch** ist die bundesweite Kampagne der gesetzlichen Unfallversicherung in Deutschland. Sie will Unternehmen und Bildungseinrichtungen dabei unterstützen eine Präventionskultur zu entwickeln, in der Sicherheit und Gesundheit Grundlage allen Handelns sind. Weitere Informationen unter www.kommmitmensch.de

Impressum

Herausgegeben von:

Deutsche Gesetzliche
Unfallversicherung e.V. (DGUV)

Glinkastraße 40
10117 Berlin
Telefon: 030 13001-0 (Zentrale)
Fax: 030 13001-9876
E-Mail: info@dguv.de
Internet: www.dguv.de

Autorinnen und Autoren:

Projektleitung:

Hauke, Angelika (DGUV/IFA)
Schwinger-Butz, Ellen (BGN)
Wetzstein, Dr. Annekatriin (DGUV/IAG)

Projektgruppe:

Hettrich, Dr. Ralph (BG ETEM)
Kallache, Dr. Malaak (BG BAU)
Mischke, Dr. Monika (BG Verkehr)
Olschok, Julia (VBG)
Roth, Sigrid (BGHW)

Unter Mitarbeit von:

Beck, Wolfgang (DGUV/SIGE)
Fuhrmann, Hans-Jochem (BG BAU)
Kromark, Kathrin (BGW)
Nold, Dr. Helmut (BG RCI)
Otto, Silke (BGHM)
Stiller-Wüsten, Claudia (BGW)
Wiegatz, Corinna (UK NRW)

Sachgebiet „Evaluation“ des Fachbereichs „Organisation von Sicherheit und Gesundheit (FB ORG)“ der DGUV.

Ausgabe: Januar 2020

DGUV Information 211-043
zu beziehen bei Ihrem zuständigen Unfallversicherungsträger oder
unter www.dguv.de/publikationen Webcode: 211043

Bildnachweis

Abb. 4: ©bloomline/stock.adobe.com; Abb. S. 48: ©Created by Freepik;
Abb. S. 49: ©Péter Mács-Fotolia

Gute Praxis der Evaluation von Präventionsmaßnahmen in der gesetzlichen Unfallversicherung

Inhalt

	Seite
1	Präambel 5
2	Definitionen 6
3	Einordnung des Evaluationsbegriffs im Kontext anderer Begriffe und Konzepte 7
4	Typische Evaluationsgegenstände im Kontext der gesetzlichen Unfallversicherung 10
5	Arten von Evaluationen 12
6	Vergleichsmaßstäbe in Evaluationen 16
7	Häufig genutzte Evaluationsmodelle 17
8	Ablauf einer Evaluation:
	Von der Planung über die Durchführung zu den Ergebnissen 23
8.1	Den Gegenstand (Maßnahme, Produkt, Projekt) der Evaluation verstehen 24
8.2	Ziele der Maßnahme, des Produkts oder Projekts verstehen 25
8.3	Zielgruppe/Betroffene der Maßnahme, des Produkts oder Projekts 25
8.4	Beteiligte an der Evaluation 26
8.5	Zweck/Hintergrund der Evaluation 28
8.6	Ermittlung von Fragestellungen der Evaluation, logischem Modell und Indikatoren 29
8.6.1	Fragestellungen der Evaluation 29
8.6.2	Das logische Modell (Wirkungsmodell) 29
8.6.3	Indikatoren 32
8.7	Modalitäten der Durchführung: Design und Methodik 34
8.7.2	Methoden und Instrumente in der Evaluation 35
8.8	Datenerhebung 38
8.9	Auswertung der Daten 38
8.10	Ergebnisdarstellung und Präsentation 38
9	Herausforderungen und Fallstricke bei Evaluationen 39
10	Fazit 43
11	Anlage 44
	Hilfsmittel zur Durchführung von Evaluationen 44
	Literatur 50
	Weiterführende Literatur der DGUV 53
	Abbildungs- und Tabellenverzeichnis 53

1 Präambel

Ein zentrales Element des gesetzlichen Auftrags der Unfallversicherung ist die Prävention. Denn „... Arbeitsunfälle und Berufskrankheiten sowie arbeitsbedingte Gesundheitsgefahren ...“ sind „... mit allen geeigneten Mitteln ...“ zu verhüten^[1].

Daher entwickelt die gesetzliche Unfallversicherung vielfältige Präventionsangebote für die Unternehmensleitungen, Führungskräfte sowie Versicherten ihrer Mitgliedsunternehmen und -einrichtungen. Ziel dieser Angebote ist die Sicherheit und Gesundheit der Beschäftigten bei der Arbeit.

Entscheidend ist für eine systematische betriebliche und überbetriebliche Präventionsarbeit, dass Präventionsmaßnahmen wirksam sind. Sie sollen die Zielgruppen erreichen und von diesen erfolgreich angewandt, genutzt bzw. umgesetzt werden. Die Überprüfung der Nutzung und Wirksamkeit von Präventionsmaßnahmen erfolgt durch Evaluationen.

Die DGUV Information 211-043 vertieft den DGUV Grundsatz 311-001 eines gemeinsamen Verständnisses zur Evaluation von Präventionsmaßnahmen in der gesetzlichen Unfallversicherung. Sie richtet sich an die Beschäftigten, die mit Evaluationsaufgaben bei den Unfallversicherungsträgern betraut sind, sowie an alle, die sich für Evaluation von Präventionsmaßnahmen interessieren und dient insbesondere dazu,

- den Begriff Evaluation einheitlich zu verwenden und im Kontext inhaltlich ähnlicher Begriffe einzuordnen,
- ein gemeinsames Verständnis weiterer zentraler Begriffe im Rahmen von Evaluationen zu etablieren,
- Ziele und Zwecke von Evaluationen festzulegen und zu beschreiben,
- die detaillierte Planung und Durchführung von Evaluationen nach einem standardisierten Ablauf zu unterstützen und
- eine Evaluationskultur in den Unfallversicherungsträgern und ihren Mitgliedsunternehmen zu verankern.

2 Definitionen

Evaluation im allgemeinen Sprachgebrauch bedeutet die sachgerechte Bewertung oder Beurteilung von Sachverhalten, Prozessen, Gegenständen oder Organisationseinheiten.

Evaluation im wissenschaftlichen Sinne ist die „explizite und systematische Verwendung [sozial]-wissenschaftlicher Forschungsmethoden zur Beschreibung und Bewertung bestimmter Gegenstände oder Maßnahmen hinsichtlich Zielsetzung und Planung, Einrichtung und Durchführung sowie Wirksamkeit und Effizienz“^[2].

Erläuterung:

Evaluation in der Prävention umfasst somit nicht nur die systematische Untersuchung der Wirksamkeit und Effizienz einer Präventionsmaßnahme (z. B. Programme, Projekte, Leistungen, Rechtsvorschriften). Evaluation kann auch die Maßnahmenentwicklung, -implementierung und -durchführung in den Fokus nehmen und somit eine vorausschauende oder begleitende Beurteilung und Weiterentwicklung der zu evaluierenden Präventionsmaßnahme umfassen. Evaluationen basieren auf einem Evaluationskonzept und auf der Anwendung wissenschaftlicher Forschungsmethoden. Daher gehen sie über simple Bewertungen per Smiley oder Sternchen hinaus: Die erzielten Ergebnisse und Schlussfolgerungen einer Evaluation beruhen auf mit wissenschaftlichen Methoden empirisch gewonnenen, beobachteten und/oder gemessenen Daten, sind nachvollziehbar sowie reproduzierbar und werden anhand zuvor festgelegter messbarer Ziele bewertet.

Evaluationen haben eine klare Funktion. Sie beantworten vorher festgelegte Fragestellungen, sind methodenoffen, können quantitative und qualitative Aussagen treffen und intendierte und nicht-intendierte Wirkungen messen.

Hinweis:

Evaluationen von Präventionsmaßnahmen der gesetzlichen Unfallversicherung und ihrer Mitgliedsunternehmen sollen möglichst von Expertinnen und Experten durchgeführt und/oder begleitet werden, die

- in der Planung, Implementierung, Durchführung und Evaluation betrieblicher und überbetrieblicher Präventionsmaßnahmen erfahren sind,
- sich kontinuierlich zu Evaluationsmethodik und Evaluationsforschung weiterbilden,
- mit deskriptiven und analytischen statistischen Verfahren vertraut sind und
- über Grundlagenwissen zu Sicherheit und Gesundheit verfügen.

Diese Evaluatorinnen und Evaluatoren sind für den kompletten Ablauf der Evaluation verantwortlich – von der Klärung der Ziele der Evaluation über die Erstellung eines Evaluationskonzepts und dessen Umsetzung bis hin zur Bewertung und Darstellung der Ergebnisse.

3 Einordnung des Evaluationsbegriffs im Kontext anderer Begriffe und Konzepte

Der Begriff „Evaluation“ hat einen engen Bezug zu oder sogar Überschneidungen mit anderen Begriffen, die oft fälschlicherweise als Synonym für Evaluation verwendet werden. Dazu zählen z. B. Controlling, Benchmarking, Qualitätssicherung, (Lern-)erfolgskontrolle, Rückmeldung, Befragung, Statistik, Analyse, Test und Bewertung. In diesem Kapitel sollen diese Begriffe kurz vorgestellt und ihr Bezug zur Evaluation geklärt werden.

Begriffe zur Bezeichnung von Methoden

Gängige Methoden, die auch bei Evaluationsvorhaben eingesetzt werden können, sind z. B. Befragung, Umfrage, Rückmeldung/Feedback, Lernerfolgskontrolle und Monitoring.

Viele dieser Methoden können als **Methoden zur Wirksamkeitskontrolle** eingesetzt werden. D. h., mithilfe der Methoden soll herausgefunden werden, ob die Maßnahme die Zielgruppe erreicht und dort die gewünschte Wirkung erzielt hat. Ein Beispiel ist die am Ende eines Seminars eingesetzte Lernerfolgskontrolle, mit der überprüft wird, ob die Teilnehmenden die verschiedenen Inhalte behalten haben.

Monitoring ist ein übergeordneter Begriff für verschiedene Arten der Überwachung von Systemen, Zuständen oder Vorgängen durch systematisches Erfassen, Protokollieren, Messen oder Beobachten. Durch Monitoring kann ermittelt werden, ob ein bestimmter Prozess wie gewünscht und geplant verläuft oder ob ggf. nachgesteuert werden muss. Monitoring und Dokumentation sind in der Prozessevaluation wesentlich^[3]. Während Monitoring Hinweise darauf gibt, dass etwas nicht nach Plan läuft, kann die Evaluation Hinweise darauf geben, warum das so ist^[4].

Begriffe wie z. B. Statistik oder Test bezeichnen **Methoden zur Datenanalyse** in Evaluationen.

Begriffe zur Beschreibung von Instrumenten zur Vorbereitung zielsetzungsgerechter Entscheidungen

Auch Ergebnisse der folgenden Instrumente können u. U. in Evaluationen genutzt werden:

Controlling ist „ein Teilbereich des unternehmerischen Führungssystems, dessen Hauptaufgabe die Planung, Steuerung und Kontrolle aller Unternehmensbereiche ist. Im Controlling laufen die Daten des Rechnungswesen[s] und anderer Quellen zusammen“^[5]. Konkret handelt es sich dabei um Kosten und Erlöse. Anhand dieser Informationen sollen die richtigen unternehmerischen Entscheidungen getroffen werden. Kosten und Erlöse können bei Evaluationen auch betrachtet werden, dienen aber nicht als alleinige Entscheidungsgrundlage.

Die **Balanced Scorecard (BSC)** dient als wichtiges Instrument der Unternehmenssteuerung: Sie setzt eine Unternehmensvision bzw. -strategie in entsprechende Ziele und Kennzahlen um und orientiert sich bei der Beurteilung der Umsetzung der Strategie ausschließlich an diesen Kennzahlen. Zur Umsetzung der Vision mittels einer Strategie werden kritische Erfolgsfaktoren aus vier verschiedenen Perspektiven heraus identifiziert^[6]:

- Finanzwirtschaftliche Perspektive (Steigerung des Unternehmenswerts, Renditen, kurzfristige Gewinne etc.)
- Kundenperspektive (z. B. Kundenzufriedenheit, Kundenakquise, Gewinn von Marktanteilen in Marktsegmenten)
- Interne Perspektive (kritische Prozesse, in denen das Unternehmen Verbesserungsschwerpunkte sieht)
- Lern- und Entwicklungsperspektive (Identifikation von Infrastrukturen für langfristigen Erfolg des Unternehmens)

Dann werden Kennzahlen für jeden Erfolgsfaktor festgelegt, anhand derer die Umsetzung der Unternehmensstrategie bzw. die Zielerreichung gemessen werden kann. So werden bei diesem Management- bzw. Steuerungskonzept neben finanziellen auch andere Aspekte betrachtet. Auch werden Aktionsprogramme entwickelt, die die Zielerreichung fördern sollen. Kennzahlen können auch in quantitativen Evaluationen berücksichtigt werden.

Benchmarking ist ein Instrument der Wettbewerbsanalyse. Es handelt sich dabei um einen kontinuierlichen Vergleich mit (mehreren) Unternehmen hinsichtlich Produkten, Dienstleistungen, Prozessen oder Methoden^[7]. Häufig werden im Vergleichsprozess die Dimensionen Qualität, Zeit und Kosten betrachtet. Deren Gewichtung kann unterschiedlich ausgeprägt sein und sich im Extremfall alleine auf die Kostenreduzierung beschränken. Ziel ist es, die Leistungslücke zum sogenannten Klassenbesten (Unternehmen, die Prozesse, Methoden etc. hervorragend beherrschen) systematisch zu schließen. Es geht darum, voneinander zu lernen, sich an „den Besten“ zu orientieren und Handlungsempfehlungen zur Prozessoptimierung anhand der folgenden Fragen zu erarbeiten: Welche Unterschiede bestehen? Warum bestehen sie? Welche Verbesserungsmöglichkeiten gibt es?

Es handelt sich dabei eher um eine strukturierende Methode, anhand derer weitere Maßnahmen abgeleitet werden können. Eine Bewertung der Maßnahmen erfolgt i. d. R. weniger. Die gesetzliche Unfallversicherung vergleicht beispielweise trägerübergreifend die umgesetzten Qualifizierungsprozesse und erarbeitet anhand von Kennzahlen in den Dimensionen Qualität, Zeit und Kosten Handlungsempfehlungen für effektive und wirtschaftliche Verfahren^[8].

Begriffe aus der Qualitätssicherung

Qualitätssicherung umfasst „als Bestandteil des Qualitätsmanagements alle organisatorischen und technischen Maßnahmen, die vorbereitend, begleitend und prüfend der Schaffung und Erhaltung einer definierten Qualität eines Produkts oder einer Dienstleistung dienen“^[9]. Hier können Evaluationsergebnisse mit einfließen bzw. herangezogen werden.

Audit (auch Auditierung) ist der angloamerikanische Fachausdruck für „Revision“^[10]. Es handelt sich um einen Begutachtungsprozess durch von dem zu prüfenden Verantwortungsbereich unabhängige Personen. Die gesetzliche Unfallversicherung ist häufig an der Auditierung von Arbeitsschutzmanagement-Systemen von Unternehmen beteiligt. Am Ende eines Audits steht oft eine Zertifizierung, aber auch die Prüfung sowie Vergabe eines Gütesiegels. Die Anforderungen (z. B. vertragsgemäße Vereinbarungen, Standards), die erfüllt sein müssen, werden dabei im Vorfeld genau definiert. Audits sind fester Bestandteil von Qualitätsmanagementsystemen.

Gemeinsam ist den Instrumenten zur Vorbereitung zielsetzungsgerechter Entscheidungen eine starke Orientierung an Indikatoren bzw. Kennzahlen. Die Verwendung von Kennzahlen ermöglicht oftmals eine quantitative Bewertung.

Evaluationen betrachten sehr viel stärker den Weg zur Zielerreichung. Sie überprüfen, ausgehend von einem logischen Modell (Wirkungsmodell), die Mechanismen der Zielerreichung. Dazu nutzen Evaluationen qualitative und/oder quantitative Methoden (vgl. Kap. 8.7.2).

Überschneidungen zwischen Evaluations- und Forschungsaktivitäten

Ein weiterer Begriff, der oft in Zusammenhang mit dem Begriff Evaluation verwendet wird, ist **Forschungsprojekt**. Eine Evaluation kann, muss aber nicht, Bestandteil eines Forschungsprojekts sein. Ob ein Forschungsprojekt eine Evaluation umfasst, hängt von dessen Zielsetzung ab. I. d. R. zielt ein Forschungsprojekt darauf ab, neues Wissen oder neue Methoden zu generieren und die Gültigkeit von Erkenntnissen für und die Anwendbarkeit von Methoden auf neue Zielgruppen zu überprüfen. Bezieht sich ein Forschungsprojekt auf die Entwicklung einer Maßnahme in Bezug auf ein definiertes Ziel, so umfasst ein Forschungsprojekt i. d. R. den Nachweis der theoretischen Wirksamkeit dieser Maßnahme hinsichtlich des Ziels (z. B. reduziert der Einsatz von staubarmen Mehlen allergische Atemwegserkrankungen?). Eine Evaluation hingegen prüft, ob eine theoretisch erfolgreiche Maßnahme auch praktisch erfolgreich ist, indem sie die Zielgruppe erreicht und dort erfolgreich angewandt wird. In der Praxis bedeutet dies am Beispiel: „Ist das Mehl auf dem Markt verfügbar?“ „Wird es gekauft und richtig angewendet?“ „Erreicht es das Ziel?“ In einem Forschungsprojekt geht es also primär darum festzustellen, ob eine Maßnahme theoretisch wirksam ist oder wie sie verändert werden muss, um wirksam zu sein. In einer Evaluation geht es hingegen um die Frage, ob eine theoretisch wirksame Maßnahme auch praktisch ihre Wirkung entfaltet.

Abgrenzung von Evaluations- gegenüber Aufsichtstätigkeiten

Auch der Begriff der **Aufsichtstätigkeit** wird im Kontext der gesetzlichen Unfallversicherung mit dem Begriff der Evaluation in Zusammenhang gebracht. Die Aufsichtspersonen der Unfallversicherungsträger haben nach §17 SGB VII den hoheitlichen Auftrag, die Mitgliedsunternehmen zu überwachen und zu beraten^[1]. Der umfassende Überwachungsauftrag bezieht sich auf das gesamte Aufgabenspektrum der Prävention und erfasst somit alles Handeln und Unterlassen, welches mit der Prävention in Zusammenhang steht. Übergeordnetes Ziel von Betriebsbesichtigungen ist, sich vor Ort ein Bild von Sicherheit und Gesundheit bei der Arbeit zu machen. Unternehmen werden hinsichtlich möglicher Gefährdungen und Belastungen überprüft, die zu Arbeitsunfällen, Berufskrankheiten und anderen arbeitsbedingten Erkrankungen führen können. Als Bewertungsgrundlage dienen zahlreiche Regelwerksvorgaben. Es wird insbesondere überprüft, ob und in welchem Umfang die verbindlichen Vorgaben eingehalten bzw. umgesetzt werden. Zugleich werden die versicherten Unternehmerinnen und Unternehmer zu diesen sowie weiteren Themen beraten und somit aktiv bei der Umsetzung ihrer vielfältigen rechtlichen Verpflichtungen unterstützt. Mit diesem Vorgehen soll letztlich eine dauerhafte Gefährdungsbeseitigung oder zumindest eine Minimierung des daraus resultierenden Restrisikos erzielt werden. Die technische Prüfung fokussiert sich auf die sicherheitstechnische Prüfung von Maschinen und Geräten. Wird sie vom Aufsichtsdienst ausgeführt, ist sie eine Teilaufgabe im Rahmen der Aufsichtstätigkeit. Die Aufsichtstätigkeit ist demnach keine Evaluation, sondern eher vergleichbar mit einem Audit, bei dem eine Überprüfung der betrieblichen Situation durch eine vom Betrieb unabhängige Person erfolgt.

4 Typische Evaluationsgegenstände im Kontext der gesetzlichen Unfallversicherung

Der gesetzliche Auftrag der Unfallversicherung ist es, mit allen geeigneten Mitteln Arbeitsunfälle, Berufskrankheiten sowie arbeitsbedingte Gesundheitsgefahren zu verhüten und für eine wirksame Erste Hilfe zu sorgen^[1].

Daher bieten die Unfallversicherungsträger ihren Mitgliedsunternehmen und Versicherten ein weites Spektrum vielfältigster Maßnahmen und Angebote im Bereich der Prävention an. Um die Ziele von Sicherheit und Gesundheit bei der Arbeit und auf dem Arbeitsweg für möglichst viele Versicherte zu erreichen, werden träger- und branchenspezifische Angebote und Maßnahmen entwickelt, die den zehn „Präventionsleistungen der Unfallversicherungsträger“^[11] zugeordnet werden können (vgl. Abbildung 1).

Besonders wichtig ist angesichts begrenzter personeller und finanzieller Ressourcen der Unternehmen, Dienstleistenden und Unfallversicherungsträger der Nachweis über die Wirksamkeit von Präventionsmaßnahmen. Evaluation dient in diesem Kontext als zentrales Instrument zur Steuerung und Qualitätssicherung.

Evaluationen in der gesetzlichen Unfallversicherung haben nicht eine dieser zehn Präventionsleistungen selbst zum Gegenstand, sondern zielen immer auf die den Präventionsleistungen zugeordneten konkreten Präventionsangebote und -maßnahmen. So können zum Beispiel für die Präventionsleistung „Information, Kommunikation

und Präventionskampagnen“ Kampagnen, Filme, Arbeitshilfen und Broschüren evaluiert werden.

In der Prävention werden sowohl einzelne als auch komplexe Maßnahmen evaluiert, um deren Zielerreichung zu prüfen. Welche spezifischen Ziele im Einzelnen mit einer Maßnahme angestrebt werden, z. B. Bekanntheitsgrad, Relevanz für betriebliche Aktivitäten, Verhaltensänderung, kann je nach Unfallversicherungsträger, Branche und Anlass für die Projektierung einer solchen Maßnahme unterschiedlich sein.

Im Folgenden werden die einzelnen Präventionsleistungen kurz erläutert.

Tabelle 1 führt beispielhaft für jede Präventionsleistung mögliche Präventionsangebote und -maßnahmen auf. Mithilfe von Evaluationen kann geprüft werden, ob die jeweilige Maßnahme ihre vorgegebenen spezifischen Ziele erreicht hat. In der Regel sollen Antworten auf folgende Fragen gefunden werden:

- Wird die Maßnahme/das Produkt von der Zielgruppe genutzt?
- Wie wird die Maßnahme/das Produkt von der Zielgruppe bewertet? Wie wird es akzeptiert? Ist die Zielgruppe zufrieden?
- Hat die Zielgruppe durch die Maßnahme/das Produkt ihr Wissen/ihre Kenntnisse erweitert?

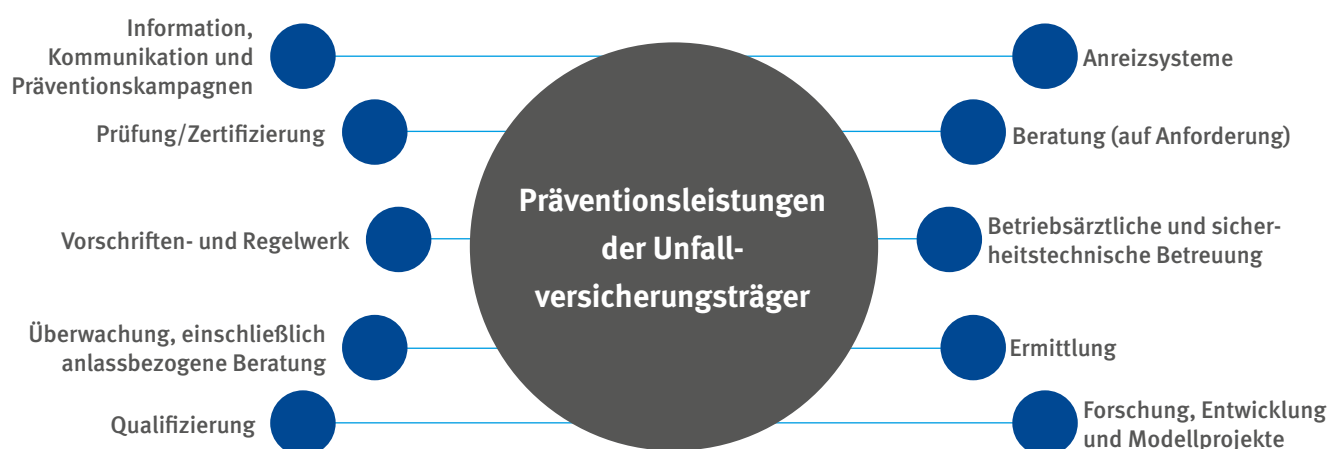


Abb. 1 Präventionsleistungen der Unfallversicherungsträger (Vergleich Quelle: Präventionsleistungen der Unfallversicherungsträger der Deutschen Gesetzlichen Unfallversicherung, Stand Dezember 2016 – aktualisiert Mai 2019. Hrsg.: Deutsche Gesetzliche Unfallversicherung e. V. (DGUV), Berlin 2016 <https://publikationen.dguv.de/dguv/pdf/10002/12471.pdf>)

- Welche Veränderungen bezogen auf das Verhalten der Zielgruppe konnten erreicht werden?
- Welche Auswirkungen entstehen dadurch in den Mitgliedsunternehmen? Wie verändern sich die Verhältnisse in den Mitgliedsunternehmen? Welche Auswirkungen hat die Maßnahme/das Produkt auf betriebliche Kennzahlen?

Tabelle 1: Präventionsleistungen und mögliche Evaluationsgegenstände

Präventionsleistung	Beispiele möglicher Evaluationsgegenstände
Anreizsysteme	Branchenspezifische Prämienverfahren, Preisvergabe für Best Practices, Gütesiegel der UVT
Beratung (auf Anforderung)	Beratung z. B. zu/zur <ul style="list-style-type: none"> • Organisation von Sicherheit und Gesundheit im Betrieb • Arbeitsschutzmanagementsystemen • Einführung neuer Arbeitsmittel • Verbesserung der Arbeitsplatzergonomie • betrieblichem Gesundheitsmanagement
Betriebsärztliche und sicherheitstechnische Betreuung	Betriebsärztliche und sicherheitstechnische Betreuung von KMU durch Kompetenzzentren, Arbeitsmedizinische Dienste der UVT
Ermittlung	Ermittlung nach Arbeitsunfällen, Berufskrankheiten, §3-Maßnahmen (Hautschutz), statistische Auswertungen
Forschung, Entwicklung und Modellprojekte	Branchenspezifische Präventionsforschung zum Erkenntnisgewinn, einschließlich Indikatoren- und Instrumentenentwicklung, wie Datenerhebung durch Messreihen, Konzeptentwicklung einschließlich Maßnahmenableitung, Pilotprojekte neuer Präventionsmaßnahmen und -produkte
Information, Kommunikation und Präventionskampagnen	Zielgruppenspezifische Medienangebote (diverse Printmedien, elektronische Medien, Medienpakete), Kampagnen, Internet, Veranstaltungen, Messeauftritte
Prüfung/Zertifizierung	Prüfgrundsätze, produkt- und personenbezogene Zertifizierungen, Arbeitsschutzmanagementsysteme (AMS), wie von einigen UVT angeboten
Vorschriften- und Regelwerk	DGUV-Vorschriften (z. B. DGUV Vorschrift 2 „Betriebsärzte und Fachkräfte für Arbeitssicherheit“), DGUV Regeln oder DGUV Informationen
Qualifizierung	Zielgruppenspezifische Aus- und Fortbildungen
Überwachung, einschließlich anlassbezogene Beratung	Überwachungstätigkeit, Umsetzung der empfohlenen Maßnahmen aus der anlassbezogenen Beratung

5 Arten von Evaluationen

In Anlehnung an Bortz und Döring (2016) werden verschiedene Arten von Evaluationen unterschieden^[12]. Die Autoren schlagen dabei folgende Einteilungskriterien vor:

- die Rollenverteilung zwischen den Evaluierenden und den Evaluierten
- die Funktionen von Evaluationen
- die Aspekte eines Evaluationsgegenstandes, die zu einem bestimmten Zeitpunkt untersucht werden und
- die in der jeweiligen Evaluation genutzte Art der Datenerhebung und -auswertung.

Die Einteilung bietet eine Hilfestellung, welche Arten von Evaluationen in einer konkreten Evaluation zum Einsatz kommen. Denn i. d. R. werden mehrere der nachfolgend genannten Arten von Evaluationen kombiniert. Welche Arten von Evaluationen angewandt werden können, hängt auch vom Zeitpunkt ab, zu dem die Evaluatoreninnen und Evaluatoren in ein Projekt eingebunden werden (bereits zur Maßnahmenentwicklung, vor Implementierung der Maßnahme, während der Maßnahmendurchführung oder erst danach).

Erläuterungen zu den folgenden Begrifflichkeiten finden sich ebenfalls im Glossar der Gesellschaft für Evaluation (DeGEval)^[13].

Evaluationsarten hinsichtlich der Rolle der Evaluierenden

Interne und externe Evaluation

Abhängig von der Rolle der Evaluatoreninnen und Evaluatoren lassen sich interne von externen Evaluationen unterscheiden. Wird die Präventionsmaßnahme von Personen evaluiert, die aus der gleichen Einrichtung kommen wie diejenigen, die die Präventionsmaßnahme gestaltet und durchgeführt haben, handelt es sich um eine interne Evaluation. Beispielsweise setzt das Kampagnenteam der DGUV eine Kampagne um und das Evaluationsteam der DGUV evaluiert diese Kampagne.

Eine spezifische Form der internen Evaluation ist die Selbstevaluation. Selbstevaluation liegt vor, wenn die Bewertung von Personen vorgenommen wird, die auch selbst an der Gestaltung der Präventionsmaßnahme beteiligt sind.

Von externer Evaluation (Fremdevaluation) wird gesprochen, wenn Expertinnen und Experten mit der Beschreibung und Bewertung beauftragt sind, die nicht zur Firma oder Einrichtung gehören. Der Evaluationsauftrag wird dann z. B. an einen Dienstleistenden außerhalb des eigenen Unfallversicherungsträgers gegeben.

Externe Evaluierende haben Expertise in Evaluationen, oft mehr Erfahrung aus dem Vergleich mit anderen, sind neutral und bringen neue Sichtweisen ein. Interne Evaluierende kennen den eigenen strukturellen Aufbau, die Präventionsthemen und -maßnahmen sowie die Zielgruppen besser und erkennen so leichter wichtige Aspekte im Detail.

Die Entscheidung, ob eine Evaluation intern oder extern durchgeführt wird, kann politisch begründet sein. In manchen Fällen kann beispielsweise die Akzeptanz bei der Zielgruppe oder den Auftraggebenden größer sein, wenn die Evaluation extern durchgeführt wird.

Entscheidend für die Wahl zwischen interner und externer Evaluation ist immer auch der Umfang der geplanten Evaluation sowie die vorhandene sozialwissenschaftliche Expertise im eigenen Haus. Ist der Umfang groß und kaum oder keine sozialwissenschaftliche Expertise vorhanden, sollte die Evaluation unbedingt durch eine externe Stelle durchgeführt werden. Kleinere Evaluationen können häufig auch intern und/oder mit Begleitung einer externen Stelle durchgeführt werden, insbesondere dann, wenn das entsprechende Know-how vorhanden ist.

Evaluationsarten hinsichtlich der Funktion von Evaluationen

Formative Evaluation und summative Evaluation

Eine formative Evaluation erfüllt eine Optimierungsfunktion und erfolgt prozessbegleitend zur Entwicklung und Implementierung einer Maßnahme. Durch diese begleitende Evaluation werden die Entwicklung und Durchführung der Maßnahme sowie deren Wirkungen fortlaufend kontrolliert und dokumentiert. Es werden Stärken und Schwächen identifiziert und wenn nötig Veränderungen initiiert. Die formative Evaluation wird zur kontinuierlichen Verbesserung eines Maßnahmenkonzepts und seiner Umsetzung angewandt sowie zur Identifikation von Wirkungsketten. Der Fokus liegt auf der prozessbegleitenden Gestaltung. Die formative Evaluation bedient sich häufig qualitativer Methoden, um Änderungsansätze, -möglichkeiten und -wünsche zu identifizieren^[12].

Die summative Evaluation erfüllt eine Kontroll- und Legitimationsfunktion. Es geht darum, eine abschließende, zusammenfassende Bilanz zu einem Evaluationsgegenstand zu ziehen. Dabei kann z. B. die Wirksamkeit (Effektivität) einer Präventionsmaßnahme bestimmt werden (Wurden die Ziele der Maßnahme erreicht?) und/oder die Frage untersucht werden, ob der Aufwand der Präventionsmaßnahme im Verhältnis zur Zielerreichung steht (Effizienz). Idealerweise erfolgt die Bilanzierung durch den Vergleich eines Zustands *vor* der Maßnahme und *nach* der Maßnahme. So kann die summative Evaluation auch eine Entscheidungsfunktion haben, nämlich dann, wenn ihre Ergebnisse entscheidend dafür sind, ob die Maßnahme fortgesetzt oder beendet wird.

Evaluationsarten hinsichtlich dessen, welche Aspekte des Evaluationsgegenstandes zu welchem Zeitpunkt untersucht werden

Konzeptevaluation

Bei einer Konzeptevaluation wird das Konzept einer Präventionsmaßnahme vor ihrer flächendeckenden Einführung beurteilt, z. B. der Plan oder das Modell einer Präventionsmaßnahme, der/das beschreibt, wie sie aufgebaut ist und wie sie idealerweise umgesetzt werden sollte (z. B. das Konzept einer neuen Broschüre).

Prozessevaluation (auch Durchführungs- oder Implementierungsevaluation)

Die Prozessevaluation ermöglicht die Erfassung des Kontexts einer Präventionsmaßnahme. Der Kontext meint dabei Faktoren im Umfeld der Maßnahme, die die Implementierung und Wirkung beeinflussen können, beispielsweise politische, rechtliche, organisationale, soziale oder kulturelle Aspekte oder Ereignisse. Die Prozessevaluation umfasst auch die Dokumentation der Umsetzungsgenauigkeit bzw. -treue eines Programms (Fidelity). In Anlehnung an Century, Rudnick und Freeman (2010) ist Fidelity definiert als „The extent to which the critical components of an intended program are present when that program is enacted.“ (Das Ausmaß der Anwesenheit kritischer/zentraler Komponenten eines beabsichtigten Programms, wenn das Programm umgesetzt wird.)^[14]. Eine Prozessevaluation ist auch dann nötig, wenn überprüft werden soll, wie/über welche Schritte/Wirkungskette eine Maßnahme wirkt (Wirkungsevaluation). Mithilfe einer Wirkungsevaluation können nicht nur die impliziten Wirkmechanismen der evaluierten Maßnahme analysiert und offengelegt werden. Es kann auch untersucht werden, welche u. U. nicht-intendierten Wirkungen die Implementierung des Evaluationsgegenstandes zur Folge hat^[12].

Ergebnisevaluation

Bei einer Ergebnisevaluation werden die von einer Maßnahme erzielten Effekte im Hinblick auf die von ihr angestrebten Ziele und somit ihre generelle Wirksamkeit untersucht.

In der Praxis werden die Ergebnisevaluation und die Prozessevaluation häufig kombiniert^[15; 16]. Die Evaluierenden können in der Ergebnisevaluation gefundene Effekte durch die Prozessevaluation leichter interpretieren, da sie den Einfluss ggf. vorhandener Störgrößen auf die gefundenen Effekte besser abschätzen können als beim

Verzicht auf eine Prozessevaluation. Werden die Maßnahmenimplementierung und Ereignisse während der Maßnahmendurchführung nicht dokumentiert, entsteht eine „Black Box“ hinsichtlich der in der Ergebnisevaluation gefundenen Effekte. Können durch die Ergebnisevaluation keine Effekte der Maßnahme gefunden werden, könnte geschlussfolgert werden, dass die Maßnahme wirkungslos ist, obwohl sie nur inadäquat implementiert wurde (Implementierungsfehler). Genauso könnte bei gefundenen Effekten geschlossen werden, dass sie durch die Maßnahme verursacht sind, obwohl ein anderes Ereignis während der Maßnahmendurchführung dafür verantwortlich war (Theoriefehler). Die Prozessevaluation kann also falsche Schlussfolgerungen verhindern. Daher erhöht die zusätzliche Prozessevaluation gegenüber einer reinen Ergebnisevaluation die interne Validität der Evaluation. Eine Prozessevaluation gibt darüber hinaus Hinweise auf die Generalisierbarkeit einer Maßnahme: Unter welchen Bedingungen und für wen funktioniert die Maßnahme? Was waren Hindernisse, was Verstärker?^[16]. Häufig sehen die konkrete Konzeption einer Evaluation und die ihr zugrundeliegenden Fragestellungen sowohl die Betrachtung der Konzept- als auch der Umsetzungs- und Wirkungsebene eines Evaluationsgegenstandes vor^[12].

Ex-ante Evaluation und Ex-post Evaluation

Ex-ante und Ex-post sind Begriffe, die unabhängig vom Evaluationsgegenstand beschreiben, zu welchem Zeitpunkt eine Evaluation durchgeführt wird. In einer sogenannten Ex-ante Evaluation, die vor der Durchführung einer Präventionsmaßnahme auf Basis von Konzepten, Anträgen oder Plänen vorgenommen wird, können Aspekte wie Bedarf, Realisierbarkeit oder Erfolgsaussichten beurteilt werden.

Bei einer Ex-post Evaluation, wird eine Präventionsmaßnahme retrospektiv und nach Beendigung beurteilt. Dabei können allerdings auch Daten einfließen, die vor oder während der Programmdurchführung erhoben wurden. Der Fokus liegt auf der Nachhaltigkeit und den langfristigen Wirkungen einer Präventionsmaßnahme.

Evaluationsarten hinsichtlich der genutzten Art der Datenerhebung und -auswertung

Quantitative, qualitative und Mixed-Methods-Evaluationen

Qualitative Evaluationsstudien nutzen einen qualitativ-interpretativen Ansatz der Datenerhebung und -auswertung. Quantitative Evaluationsstudien nutzen einen quantitativ-statistischen Ansatz. Diese Unterscheidung hat in der Praxis jedoch kaum Relevanz. Meistens werden in Evaluationsstudien sowohl qualitative als auch quantitative Daten erhoben^[12].

Weitere Evaluationsarten

Input-, Output-, Outcome- und Impact-Evaluation

Bei Input- bzw. Strukturevaluationen werden persönliche, materielle und organisatorische Ressourcen bewertet, die den Leistungserbringern (z. B. den Unfallversicherungsträgern) zur Verfügung stehen und die in ein Programm (z. B. eine Präventionsmaßnahme) investiert werden. In der Prävention gelten die Trägerschaft, die Projektorganisation, die organisatorische und personelle Vernetzung, die gegebene Organisationsstruktur, die Qualifikation der Mitarbeitenden, der Leistungsauftrag sowie die finanziellen und personellen Ressourcen als mögliche Indikatoren.

Eine Output-Evaluation meint hingegen die Beschreibung oder Untersuchung der zählbaren Dienstleistungen und Produkte (Outputs) eines zu evaluierenden Programms, über die spezifische Wirkungen erreicht werden sollen (z. B. die Outputs/Komponenten einer Präventionskampagne).

Als Outcome-Evaluation wird eine Untersuchung der Auswirkungen (Outcomes) eines Evaluationsgegenstands auf der Ebene der Zielgruppe verstanden.

Von Impact-Evaluation spricht man, wenn eine Evaluation die Wirkung einer Maßnahme hinsichtlich betrieblich relevanter Kennzahlen oder auf gesellschaftlicher Ebene betrachtet.

Globale und analytische Evaluation

Bei einer globalen Evaluation wird der Fragestellung oder dem Evaluationsgegenstand eine globale Beurteilung zugeordnet. Ein Beispiel für eine globale Evaluation ist eine Zufriedenheitsabfrage am Ende eines Seminars.

Eine analytische Evaluation liegt vor, wenn auch Einzelkomponenten oder Einzelheiten des Evaluationsgegenstandes bewertet werden. Dabei lassen sich sowohl Komponenten (z. B. die Vorbereitung oder die Durchführung einer Maßnahme) als auch Dimensionen (z. B. die Schwierigkeit eines Seminars oder die Übertragbarkeit auf das eigene Arbeitsumfeld) herausgreifen. Die Beurteilung der einzelnen Komponenten oder Dimensionen können in einer Gesamtbewertung zusammengeführt oder einzeln für sich betrachtet werden.

Transferevaluation

Von Transferevaluation spricht man, wenn der Fokus der Evaluation auf der Begleitung oder Überprüfung des Transfers von Wissen und Können in die Praxis liegt, z. B. aus einem Seminar in die berufliche Praxis.

Meta-Evaluation

In einer Meta-Evaluation ist der Evaluationsgegenstand die Evaluation selbst. Sie kann begleitend oder nach Abschluss des Evaluationsprozesses (retrospektiv) durchgeführt werden. Eine Meta-Evaluation beurteilt z. B. die Qualität von Evaluationsprozessen und -ergebnissen. Mögliche Kriterien für diese Beurteilung sind die Standards für Evaluation, die von der Gesellschaft für Evaluation entwickelt wurden ^[vgl. 13].

Der Begriff der Meta-Evaluation sollte nicht mit dem Begriff der *Meta-Analyse* verwechselt werden. Eine Meta-Analyse (auch systematisches Review) ist eine (statistische) Zusammenfassung der Ergebnisse mehrerer Studien zu einer spezifischen Forschungsfrage.

6 Vergleichsmaßstäbe in Evaluationen

Bei der Evaluation von Präventionsmaßnahmen werden Referenzen im Sinne von Vergleichsmaßstäben herangezogen, die eine Bewertung der Präventionsmaßnahmen auf Basis der in der Evaluation erzielten Ergebnisse ermöglichen. Diese Vergleichsgrößen bieten in der Planungsphase eine Orientierung und in der Abschlussphase eine Hilfe bei der Interpretation der Ergebnisse.

In Anlehnung an Schmidt^[17] können folgende Referenzgrößen möglich sein. Diese sind nicht trennscharf, d. h., sie können einzeln oder kombiniert herangezogen werden. Die Auswahl der Referenzgrößen bestimmt die Ausrichtung der Evaluation.

Vergleich mit anderen Präventionsmaßnahmen

Evaluationsergebnisse zur Durchführung oder Wirksamkeit einer Präventionsmaßnahme werden mit Ergebnissen ähnlicher Präventionsmaßnahmen verglichen. Es wird keine absolute Bewertung durchgeführt, sondern eine relative. *Beispiel:* Die Zufriedenheit mit dem Seminar x und die Bewertung des Transfererfolgs von Seminar x werden mit den Ergebnissen zur Zufriedenheit mit dem Seminar y und dem Transfererfolg von Seminar y verglichen.

Vergleich zwischen angestrebten Zielen und Zielerreichung

Die Wirksamkeit der Präventionsmaßnahme wird daran festgemacht, ob die Ziele der Präventionsmaßnahme erreicht wurden. *Beispiel:* Erreicht eine Präventionskampagne die Kampagnenziele, die vor Beginn der Kampagne im Fachkonzept festgeschrieben wurden, gilt sie als erfolgreich.

Vergleich von Ergebnissen zu unterschiedlichen Messzeitpunkten

Es werden zu verschiedenen Messzeitpunkten, z. B. vor, während, unmittelbar nach Abschluss sowie Monate nach Abschluss der Durchführung einer Präventionsmaßnahme Daten erhoben, die dann miteinander verglichen werden können. Hier lässt sich der relationale Fortschritt ermitteln. *Beispiel:* Die Befragung zu einer neuen Ausbildung erfolgt vor Beginn der Ausbildung, zwischen den Ausbildungsmodulen, am Abschluss der Ausbildung und acht Monate nach Ausbildungsende. Es werden jeweils die gleichen Daten erhoben, sodass Veränderungen bei den Teilnehmenden, z. B. bezüglich ihrer Kompetenz und ihren Fertigkeiten, identifiziert werden können.

Vergleich mit normativ gesetzten oder auf Basis theoretischer Modelle abgeleiteter Qualitätsstandards

Hier wird ermittelt, ob eine Präventionsmaßnahme normativ gesetzte oder aus theoretischen Modellen abgeleitete Qualitätsstandards erreicht. *Beispiel:* In Vorschriften der gesetzlichen Unfallversicherung sind bestimmte Normen festgeschrieben. Mittels Evaluation kann festgestellt werden, ob diese Normen durch die Umsetzung der Vorschrift im Unternehmen erreicht werden.

7 Häufig genutzte Evaluationsmodelle

Für eine Reihe von Präventionsmaßnahmen lassen sich Evaluationsmodelle aus der Forschungsliteratur finden, die als Grundlage dienen. Sie lassen sich sowohl auf die Wirksamkeitsanalyse als auch auf die Qualitätsmessung von Präventionsmaßnahmen anwenden und können spezifisch angepasst werden. Je nach Evaluationskonzept und Evaluationsgegenstand wird auf unterschiedliche Evaluationsmodelle zurückgegriffen.

Es wird empfohlen bei Evaluationen auf der Basis eines Evaluationsmodells zu arbeiten. Evaluationsmodelle bieten einen Leitfaden zur Vorgehensweise bei der Evaluation. Sie lassen sich hinsichtlich bestimmter Aspekte mit anderen Modellen kombinieren.

Einige wenige Evaluationsmodelle, wie das CIPP-Modell von Stufflebeam^[18] und das Dynamic Integrated Evaluation Model von Thiele Schwarz, Lundmark und Hasson^[19], betrachten die Maßnahmenentwicklung und Evaluation als gemeinsamen Prozess und beginnen mit der Bedarfs- und Umfeldanalyse. Nach der Beschreibung dieser beiden Modelle schließt sich die Beschreibung weiterer häufig genutzter Evaluationsmodelle an.

CIPP-Modell von Stufflebeam^[18]

Stufflebeam sieht den Nutzen von Evaluation vor allem darin, Maßnahmen zu verbessern. Die Überprüfung der Wirksamkeit steht bei ihm weniger im Vordergrund: „evaluation’s most important purpose is not to prove, but to improve“^[18]. Das CIPP-Modell beschreibt, welche vier Phasen im Evaluationsprozess zu berücksichtigen sind:

Evaluation des Kontexts (C)

Die Kontextevaluation dient der Bedarfsanalyse: Was ist der Bedarf der Organisation? Was sollen Ziele der Maßnahme sein? Welche Rahmenbedingungen für Maßnahmen bestehen (Ressourcen und Problemlagen)? Wer ist die Zielgruppe? Ergebnis der Kontextevaluation ist ein verbessertes Verständnis der durch die Maßnahme erwarteten Ziele. Als Methode bietet sich die Stakeholderanalyse an.

Inputevaluation (I)

Die Inputevaluation dient dazu verschiedene Ablaufpläne der Durchführung und Umsetzung (Implementierung) der Maßnahme gegeneinander abzuwägen. Bei der Entscheidung für die beste Verfahrensalternative werden vorhandene Ressourcen und strukturelle Rahmenbedingungen berücksichtigt und über die Ressourcenverteilung entschieden. Es geht darum, die beste (realisierbarste, wirtschaftlichste, erfolgversprechendste) Implementierung der Maßnahme zu ermitteln.

Prozessevaluation (P)

Die Prozessevaluation überwacht und dokumentiert die Implementierung und Durchführung der Maßnahme und vergleicht sie mit dem Ablaufplan und den Zielen der Maßnahme. Die kontinuierliche Überwachung ermöglicht korrekatives Eingreifen. Die Dokumentation ist essentiell zur Interpretation der Maßnahmenergebnisse, d. h. zur Erklärung von Erfolgen und Misserfolgen der Maßnahme/ des Programms.

Produktevaluation (P)

Die Produktevaluation umfasst vier Aspekte: Die Einflussbewertung prüft, ob die Maßnahme die Zielgruppe erreicht hat und wie die Zielgruppe die Maßnahme einschätzt. Die Effektivitätsevaluation überprüft die Wirksamkeit der Maßnahme im Sinne der klassischen Wirksamkeitsprüfung. Die Nachhaltigkeitsevaluation prüft, ob der Nutzen für die Zielgruppe anhaltend ist. Die Transferevaluation prüft, welche Faktoren den Transfer in andere Bereiche begünstigen (Was waren Erfolgsfaktoren, was waren hinderliche Umstände, was hat zu Misserfolg geführt und warum?)^[18; 20].

Das CIPP-Modell hat sich auch für die Evaluation von Maßnahmen der Prävention und Gesundheitserhaltung bewährt.

Das Dynamische, Integrierte Evaluationsmodell (DIEM) von Thiele Schwarz, Lundmark und Hasson^[19]

Bei Interventionen im betrieblichen Kontext wird zunehmend Wert auf partizipative Entwicklung und Einpassung der Maßnahme in betriebliche Strukturen sowie Abläufe gelegt. Durch die enge Zusammenarbeit mit den betrieblichen Akteuren erhöht sich die Akzeptanz der und die Identifikation mit der Maßnahme. Gelingt die Integration in betriebliche Abläufe und Strukturen, so gehen von Thiele Schwarz, Lundmark und Hasson^[19] davon aus, dass der durch die Intervention angestoßene Veränderungsprozess eher kontinuierlich als episodisch verläuft. Daher schlagen sie ein Evaluationsmodell für betriebliche Kontexte vor, das den kontinuierlichen Charakter der Veränderung durch die Intervention aufgreift und bereits vor der Maßnahmenkonzeption ansetzt. Ihr Evaluationsmodell umfasst acht Schritte (siehe Abbildung 2). Die ersten vier bilden die Planungsphase; die letzten vier die Phase der Veränderung nach Maßnahmenimplementierung. Damit kommt im DIEM dem Prozess der Maßnahmenplanung und -(weiter)entwicklung (und über die Maßnahmenkontextualisierung auch dem Kontext) eine größere Rolle zu

als im später beschriebenen Evaluationszirkel. Dort wird die Maßnahmenkonzeption als der Evaluationsplanung vorgehend betrachtet.

Bestimmung betrieblicher Ziele und Ergebnisse (Schritt 1)

Der erste Schritt dient der Bestimmung von übergeordneten, meist langfristigen Zielen und Ergebnissen, die dem Betrieb/der Einrichtung wichtig sind (z. B. hinsichtlich Produktivität, Effizienz, Sicherheit etc.). Sie bilden den Rahmen und Kontext für eine Veränderung.

Bestimmung der Maßnahmenziele sowie gewünschter Ergebnisse der Maßnahme (Schritt 2)

Welche kurz- und langfristigen Ergebnisse sollen in Abhängigkeit von betrieblichen Zielen und gewünschten betrieblichen Ergebnissen durch die Maßnahme erreicht werden?

Zur Beantwortung der Fragen in Schritt 1 und 2 sollten im Betrieb/in der Einrichtung existierende Daten (z. B. aus Benchmarking, Gefährdungsbeurteilungen etc.) und Wissen genutzt werden. Auch aus Wissens- und Datenlücken lassen sich Maßnahmenziele ableiten. Hierfür gilt es, Gespräche mit allen Stakeholdern zu führen. Ein gemeinsames Verständnis über betriebliche Ziele und Ergebnisse

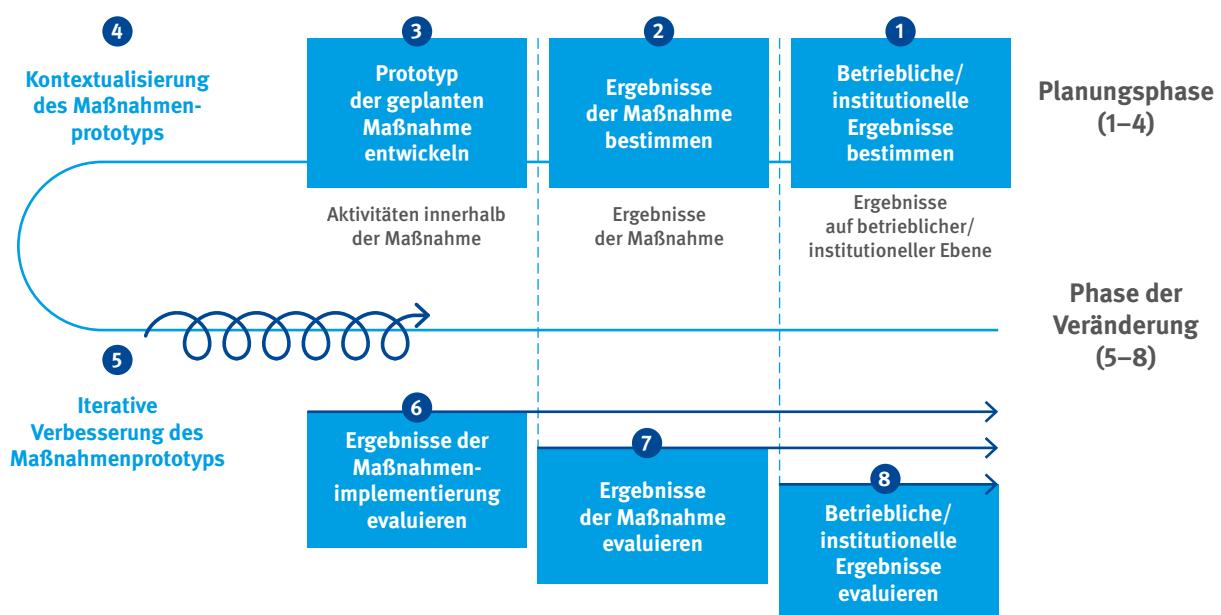


Abb. 2 Das Dynamische, Integrierte Evaluationsmodell (Quelle: von Thiele Schwarz, U.; Lundmark, R.; Hasson, H.: The Dynamic Integrated Evaluation Model (DIEM): Achieving Sustainability in Organizational Intervention through a Participatory Evaluation Approach. Stress and Health 32 (2016) Nr. 4, S. 285–293)

sowie die Maßnahmenziele muss erreicht werden. Die Maßnahmenziele sollten Arbeitgeber- und Arbeitnehmerziele enthalten, um das Commitment aller zu erhöhen.

Gemeinsame Entwicklung des geplanten Maßnahmenprototyps (Schritt 3)

Ausgehend von den gewünschten Maßnahmenzielen überlegen die Stakeholder z. B. in einem Workshop, durch welche Aktivitäten das jeweilige Maßnahmenziel erreicht werden könnte (Welche Aktivitäten sind schon realisiert? Welche müssten neu dazu kommen? Welche müssten verstärkt werden? Welche Aktivitäten sind kontraproduktiv?). Es entsteht ein logisches Modell (bzw. Wirkungsmodell). Nach der Sammlung priorisieren die Stakeholder, welche Aktivitäten den größten Einfluss haben, am besten umgesetzt werden können und die positivsten Nebeneffekte haben. Nach Einigung auf die wichtigsten Aktivitäten wird konkretisiert: Wie sieht die Intervention genau aus und wann wird sie wie oft, wo und mit wem durchgeführt?

Kontextualisierung des Maßnahmenprototyps (Schritt 4)

Der betriebliche Kontext muss vor der Implementierung der Maßnahme genau analysiert werden. Zum einen steigt dadurch die Kenntnis der unabhängigen Variablen im Evaluationsprozess. Zum anderen kann der Maßnahmenprototyp dadurch so angepasst werden, dass die Maßnahme im vorliegenden Kontext so effektiv wie möglich wirken kann. Damit erfolgen in DIEM Modifikationen hinsichtlich der in der Literatur diskutierten Fragen „what works for whom under which circumstances?“ (Was funktioniert für wen unter welchen Bedingungen?). Ungünstige Kontexte sollen als Erklärung für das Scheitern von Maßnahmen entfallen und die „Readiness for change“ (Bereitschaft für Veränderung) erhöht werden.

Konkret erfolgt die Kontextualisierung der Maßnahme mithilfe der betrieblichen Stakeholder. Ausgehend vom logischen bzw. Wirkungsmodell (vgl. Kap. 8.6.2) überlegen sie, wie jede Maßnahmenaktivität in ihrem betrieblichen Kontext konkret umgesetzt würde und wodurch Maßnahmeneffekte gehemmt/verstärkt werden könnten. So erschließt sich z. B., ob Maßnahmenaktivitäten auf anderen betrieblichen Ebenen verstärkt werden müssten, um sich durchsetzen zu können. Beispielsweise könnte eine Erkenntnis sein, dass Führungskräfte noch stärker auf die Maßnahme vorbereitet werden müssten. Konkurrierende Ziele könnten aufgedeckt und im Vorfeld angepasst werden. Das Ergebnis ist der veränderte Maßnahmenprototyp.

Iterative Verbesserung des Maßnahmenprototyps (Schritte 5 und 6)

Der veränderte Maßnahmenprototyp wird nun im Betrieb implementiert. Es bietet sich an, den Maßnahmenprototyp wie in einer Pilotstudie zunächst an einer kleinen Gruppe von Mitarbeitenden des Betriebs zu testen. Dadurch können Aussagen darüber getroffen werden, wie die Implementierung in der Zielgruppe erfolgt ist, wie die Maßnahme aufgenommen wurde und ggf. welche Effekte der Maßnahmenprototyp in der Zielgruppe schon erreicht hat. Auf Basis der Ergebnisse aus dieser ersten Implementierung wird der Maßnahmenprototyp weiter verbessert und in einer anderen Gruppe von Mitarbeitenden erneut implementiert.

DIEM zielt darauf ab, die Maßnahme durch schnelle Datenauswertungen dynamisch zu evaluieren, d. h. schnell Verbesserungsmöglichkeiten aufzudecken und die Maßnahme weiterverändert wieder zu implementieren. Ein kontinuierlicher Verbesserungsprozess soll in Gang kommen. Zur Prozessevaluation empfehlen von Thiele Schwarz, Lundmark und Hasson^[19] innovative Formen der Datenerhebung z. B. webgestützte kurze interventionsbezogene Tagebücher, kurze Selbsteinschätzungen und Beantwortungen von Fragen danach, was gut geklappt hat und was anders gemacht werden müsste sowie Beobachtung per Checkliste durch Führungskräfte.

Evaluation der Maßnahmenergebnisse und der betrieblichen Ergebnisse (Schritte 7 und 8)

Das zugrundeliegende logische Modell (bzw. Wirkungsmodell) bestimmt den/die geeigneten Zeitpunkt/e für die Evaluation der Maßnahmen- und betrieblichen Ergebnisse. D. h., das logische Modell hilft bei der Entscheidung, wann davon ausgegangen werden kann, dass die Maßnahme eine Wirkung in ihrer Zielgruppe erreicht hat, wann sich diese Wirkung in betrieblichen Ergebnissen niederschlagen konnte und sinnvollerweise evaluiert werden sollte. Zur Erfassung der betrieblichen Ergebnisse empfehlen von Thiele Schwarz, Lundmark und Hasson^[19] auf im Betrieb bereits bestehende Daten (z. B. zu Produktivität oder Qualität) zurückzugreifen. Dies erhöhe die Wahrscheinlichkeit, dass der Betrieb diese Daten auch weiter überwacht und die Maßnahme über die Zeit selbstständig weiterentwickelt.

Qualitätsmodell von Donabedian^[21]

Das Qualitätsmodell von Donabedian^[21] ist im Gesundheitswesen entstanden. Zur Strukturierung des Begriffs der „Qualität“ unterscheidet er drei Qualitätsdimensionen. Diese Dimensionen werden implizit im Exkurs des Kapitels 8.6.2 und dem dort beschriebenen logischen Modell aufgegriffen.

Die *Strukturqualität* (appraisal of structure) entspricht der Qualität der Leistungserstellung. Sie bezieht sich auf die Qualität der eingesetzten Faktoren und auf die Qualität von Aufbau- und Ablauforganisation. Sie beinhaltet persönliche, materielle und organisatorische Ressourcen, die den Leistungserbringenden zur Verfügung stehen.

Die *Prozessqualität* (assessment of process) wird gemessen an der Qualität des Dienstleistungsprozesses selbst. Sie bezieht sich auf Aktivitäten und den indikationsgerechten Einsatz der vorhandenen Strukturen bzw. Ressourcen. Im Prozess sind äußere Einflüsse möglich und zu beachten.

Die *Ergebnisqualität* (assessment of outcomes) bezieht sich auf die Qualität der Zielerreichung, also auf die Qualität des Ergebnisses. Die Beurteilung des Ergebnisses erfolgt im Hinblick auf den erreichten Zustand bei der Zielgröße (z. B. Person, Situation, System, Prozess). Die Ergebnisqualität ist der primäre Beurteilungsmaßstab für die Wirksamkeit einer Leistung. Voraussetzung für die Ergebnisqualität ist, dass Struktur- u. Prozessqualität angemessen gewährleistet sind.

Vier-Ebenen-Modell von Kirkpatrick^[22]

Es ist eines der ältesten, aber auch am häufigsten eingesetzten Evaluationsmodelle. Im Bereich der Weiterbildungsevaluation hat es sich zur Erfolgskontrolle bewährt. Das Modell beinhaltet vier Ebenen:

Reaction (= *Reaktionsebene*): Die subjektive Einschätzung der Maßnahme durch die Teilnehmenden wird als entscheidend angesehen, da angenommen wird, dass sich ihre Zufriedenheit positiv auf Motive und Interessen, Lernmotivation und damit den Lernerfolg (= Erfolg oder Misserfolg des Trainings) auswirkt.

Learning (= *Lernerfolg*): Indikatoren für den Lernerfolg sind Wissenszuwachs, verbesserte Fähigkeiten und Fertigkeiten sowie Einstellungsänderung. Ein positiver Effekt auf der Lernebene ist grundlegend für den Lerntransfer, der über die Verhaltensänderung als dritte Ebene im Modell abgebildet wird.

Behavior (= *Verhaltensebene*): Das Gelernte wird angewandt und umgesetzt, es hat nach der Einstellungs- auch eine Verhaltensänderung stattgefunden. Darauf, ob tatsächlich ein Transfer des Gelernten im Betrieb stattfindet, haben verschiedene Faktoren Einfluss. Nicht nur die individuelle Einstellungs- und damit Verhaltensänderung ist wichtig, auch die Gegebenheiten im Arbeitsumfeld spielen eine wesentliche Rolle.

Result (= *Gesamtresultat und Effektivität einer Maßnahme, betriebliche Kennzahlen*): Wurden die angestrebten Ziele erreicht? Lässt sich anhand von Kennzahlen eine Produktivitätssteigerung festmachen? Wurden Fehlzeiten gesenkt?

Nach Kirkpatrick^[22] stehen die vier genannten Ebenen in einem hierarchischen bzw. kumulativen Zusammenhang. Ohne Zufriedenheit der Teilnehmenden mit der Maßnahme gibt es keinen Lernerfolg, ohne Lernerfolg kann kein Transfer auf berufliche Situationen erfolgen, und schließlich kann es so keinen Effekt auf das Unternehmen geben. Die ersten beiden Ebenen reaction und learning sind die üblichen Schritte, die bei einer Evaluation i. d. R. erfasst werden. Es wird meist die Zufriedenheit der Teilnehmenden und der Lernerfolg ermittelt, z. B. in Form eines Wissenstests. Je höher die Ebenen werden, umso komplexer sind die Anforderungen an die Evaluation.

Informationsverarbeitungsmodell von McGuire^[23]

Dieses Modell ähnelt dem von Kirkpatrick^[22] und geht davon aus, dass zur Verarbeitung von Informationen mindestens fünf Bedingungen erfüllt sein müssen: Aufmerksamkeit, Verstehen, Akzeptieren, Beibehalten, Verhalten. Zunächst muss Aufmerksamkeit auf die Informationen gerichtet werden. Dann müssen die Informationen verstanden und die Argumente akzeptiert werden, damit es zu einer Einstellungsänderung kommt. Die geänderte Einstellung muss beibehalten und das Verhalten gemäß der neuen Einstellung angepasst werden. Das Modell macht deutlich, wie schwierig es ist, durch Informationen eine Einstellungsänderung zu bewirken. Denn schon, wenn die Zielgruppe einen der ersten Schritte nicht durchläuft bzw. nicht durchlaufen kann, ist die Kommunikation nicht erfolgreich und führt somit auch nicht zur Einstellungsänderung.

Phasenmodell der Kampagnenwirkungen der Bundeszentrale für gesundheitliche Aufklärung (BZgA)

Die Ideen des Informationsverarbeitungsmodells nach McGuire^[23] wurden auch auf die Evaluation von Kampagnen übertragen – z. B. von der BZgA zur Evaluation ihrer Präventionskampagnen^[z. B. 24; 25].

Das Phasenmodell der Kampagnenwirkungen beschreibt, dass Kampagnen ihre Wirkung über mehrere Phasen bzw. Stufen erreichen, wobei eine Phase auf der anderen aufbaut. In der Informationsphase kann über die Präsenz der Kampagne zunächst ihre Botschaft bewusst wahrgenommen und Wissen erworben werden. In der Überzeugungsphase soll sich die Einstellung entwickeln, ein neues Verhalten zu praktizieren und eine Verhaltensabsicht ausbilden (Entscheidungsphase). In der Umsetzungsphase wird das neue Verhalten praktiziert; in der Bestätigungsphase geht das neue Verhalten in Routine über und wird beibehalten.

Ebenenmodell zur Kampagnenevaluation von Wetzstein und Taskan-Karamürsel^[26]

In Abhängigkeit von den thematischen Handlungsfeldern und den Zielen bzw. Zielgruppen der Kernbotschaften bestehen Präventionskampagnen oft aus einer Kombination verschiedener Maßnahmen. Aus dieser Komplexität heraus empfiehlt sich eine Evaluation auf mehreren Ebenen. Daher haben die Autorinnen^[26] ein Ebenenmodell zur Überprüfung der Wirksamkeit von Präventionskampagnen abgeleitet. Die Ebenen einer Kampagnenevaluation werden nachfolgend erläutert und in Abbildung 3 dargestellt:

Auf der **Ebene 0** erfolgt bereits vor bzw. mit Beginn der Kampagne eine interne und externe Konzeptevaluation. Die interne Konzeptevaluation untersucht, inwiefern die Überzeugung und Akzeptanz der Kampagne bei den internen Verantwortlichen vorhanden ist, die maßgeblich an der Verbreitung der Kampagneninhalte beteiligt sind. Die externe Konzeptevaluation wird dazu genutzt, verschiedene Kampagnenmaßnahmen bereits im Vorfeld eines breiten Einsatzes an den Zielgruppen zu testen (z. B. Wie wirkt eine bestimmte Broschüre auf die Zielgruppen?).

Auf der **Ebene 1** wird der Umfang der Kampagnenaktivitäten und -maßnahmen, d. h. alle Maßnahmen aus den Bereichen Presse, Aktivitäten, Medien und Give-aways in einheitlichem Format dokumentiert. Dadurch kann eine Aussage getroffen werden, welche Bemühungen erfolgt sind, den Zielgruppen die Kampagne und ihre Botschaften zugänglich zu machen.

Auf der **Ebene 2** erfolgt anhand von qualitativen und quantitativen Merkmalen die Dokumentation und Bewertung der Medienresonanz (in Text, Bild, Radio- und TV-Sendungen). Diese Ebene gibt darüber Auskunft, wie häufig die Zielgruppen Gelegenheit hatten, aus verschiedenen Medien etwas über die Präventionskampagne zu erfahren.

Auf der **Ebene 3** wird die Wahrnehmung, Bewertung und Akzeptanz der Kampagne bei den Zielgruppen ermittelt: Haben die Zielgruppen die Kampagne tatsächlich wahrgenommen? Wird die Kampagne positiv bewertet? Werden die Themen als relevant und wichtig erachtet?

Die **4. Ebene** steigt tiefer in die Frage ein, ob sich durch die Kampagne etwas im Wissen, in den Einstellungen sowie Verhaltensweisen und den Verhältnissen der Zielgruppen geändert hat. Veränderungen auf dieser Ebene sind längerfristigen Prozessen unterworfen als auf der vorgeschalteten Wahrnehmungs-, Bewertungs- und Akzeptanzebene. Die Bewertung der vierten Ebene erfordert häufig ein komplexeres Untersuchungsdesign mit Vorher-Nachher-Messung.

Die **5. Ebene** der Auswirkungen im Betrieb soll zeigen, ob die Aktivitäten und Maßnahmen auf spezifische Kennzahlen in einem Betrieb/einer Organisation wirken können.

Die **Ebenen 6 und 7** schlagen eine Brücke zur formativen Evaluation der Kampagne und richten den Blick nach innen. Die Kampagnenverantwortlichen werden auf der 6. Evaluationsebene zur Struktur der Kampagne, zu den internen Prozessen, zur Projektarbeit und nach dem Optimierungsbedarf für die aktuelle und die nachfolgenden Kampagnen befragt. Hierzu kann im Rahmen des Evaluationsmodells auf der 7. Ebene beraten werden.

Auf der **Ebene 8** erfolgt eine Rückmeldung der Evaluatoreninnen und Evaluatoren, d. h., Evaluationsergebnisse werden strukturiert und kontinuierlich an alle relevanten Gremien zurückgemeldet.

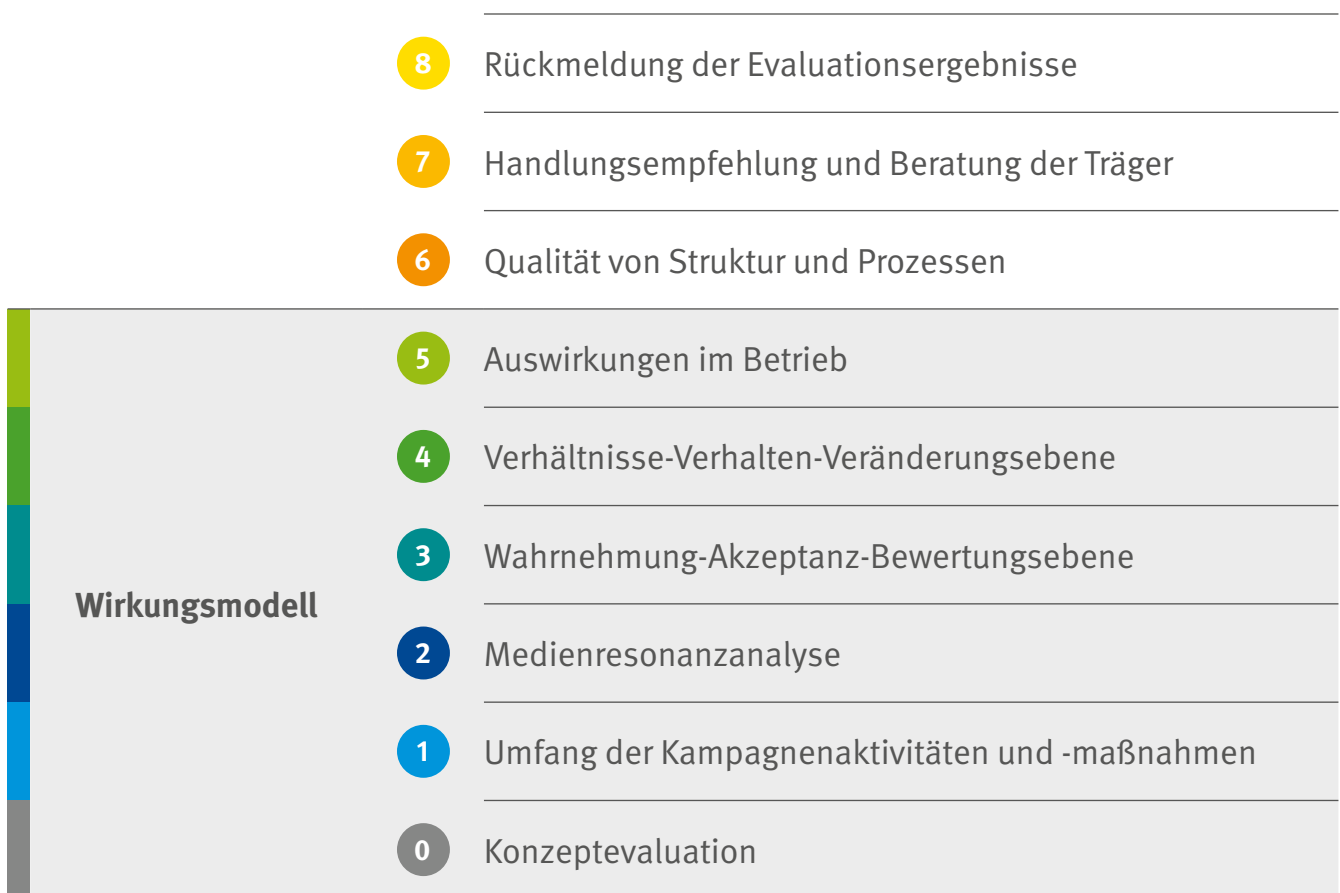


Abb. 3 Ebenenmodell der Evaluation von Präventionskampagnen (Quelle: Hessenmöller, A.; Rahnfeld, M.; Renner, S.; Wetzstein, A.: IAG Report 5/2017. Die Evaluation der Präventionskampagne „Denk an mich. Dein Rücken“ Planung, Durchführung und zentrale Ergebnisse. Hrsg.: Deutsche Gesetzliche Unfallversicherung (DGUV), Berlin 2017 https://publikationen.dguv.de/dguv/udt_dguv_main.aspx?FDOCUID=26875)^[27]

8 Ablauf einer Evaluation: Von der Planung über die Durchführung zu den Ergebnissen

Grundlage einer Evaluation ist die Evaluationsplanung anhand eines **Evaluationskonzepts**. Das Evaluationskonzept plant die komplette Evaluation vor, indem zunächst Fragen zu den folgenden zehn Punkten gestellt und im Austausch mit den Auftraggebenden beantwortet werden:

1. Maßnahme, Produkt oder Projekt, das Gegenstand der Evaluation sein soll
2. Ziele der Maßnahme, des Produkts oder Projekts
3. Zielgruppe/Betroffene der Maßnahme, des Produkts oder Projekts
4. Beteiligte an der Evaluation
5. Zweck/Hintergrund der Evaluation
6. Fragestellungen, logisches Modell und Indikatoren
7. Modalitäten der Durchführung, Design und Methodik
8. Datenerhebung
9. Auswertung der Daten
10. Ergebnisdarstellung und Präsentation

Zur Erstellung des Evaluationskonzepts ist es erforderlich, den Evaluationsgegenstand (z. B. eine Maßnahme, ein Projekt oder ein Produkt) genau kennenzulernen und zu ermitteln, welche Ziele damit erreicht, welche Veränderungen bei welcher Zielgruppe bewirkt werden sollen und was über den Evaluationsgegenstand herausgefunden werden soll. Damit wird das Erkenntnisinteresse ermittelt. Um das Evaluationskonzept zu erstellen, müssen Evaluatoreninnen und Evaluatoren umfangreiche Informationen einholen und entsprechende Fragen stellen. Dafür kann als Hilfsmittel die Checkliste zur Auftragsklärung^[28] verwendet werden.

Auf Basis der gewonnenen Informationen wird das Evaluationskonzept ausgearbeitet. Dabei werden die Fragestellungen finalisiert, es wird ein logisches Modell aufgestellt und Indikatoren abgeleitet. Des Weiteren muss festgelegt werden, mit welchen Evaluationsmethoden die Evaluation durchgeführt wird, so dass die Fragestellungen beantwortet werden können. Auch muss entschieden werden, welches Evaluationsdesign zielführend ist. Schließlich werden – soweit möglich – Festlegungen zur Datenerhebung und -auswertung getroffen sowie Art und Zeitpunkt der Ergebnispräsentation ermittelt.

Das fertige Evaluationskonzept wird – in der Regel bevor die eigentliche Durchführung der Evaluation beginnt – zwischen Auftraggeber und Evaluationsteam abgestimmt. Die Durchführung der Evaluation erfolgt auf Grundlage des Evaluationskonzepts.

Abbildung 4 stellt die zehn o. g. Prozessschritte, die im Evaluationskonzept berücksichtigt werden, als sogenannten Evaluationszirkel dar. Häufig erfolgen Evaluationen in der Praxis parallel oder nachgelagert zur eigentlichen Maßnahmendurchführung. In diesen Fällen können Evaluatoreninnen und Evaluatoren keinen Einfluss mehr auf die mit der Maßnahme verfolgten Ziele und die Zielgruppe nehmen. Dennoch sind die Informationen über diese beiden Aspekte zur Planung der Evaluation, zur Erstellung des Evaluationskonzepts und zur Auswahl/Erstellung der an die Fragestellungen der Evaluation angepassten Erhebungsmethoden wichtig.

Zusammenfassend wird deutlich, dass eine Evaluation ein komplexer Prozess ist, der deutlich über den Ansatz „wir machen mal eben einen Fragebogen und werten ihn aus“ hinausgeht. Nur wenn man den Evaluationsgegenstand, die genaue Zielgruppe und das eigentliche Erkenntnisinteresse kennt, kann man eine Evaluation mit allen Methoden so konzipieren, dass die benötigten Ergebnisse ermittelt werden können. Evaluatoreninnen und Evaluatoren ermitteln zunächst mit den Auftraggebenden, welche Erkenntnisse durch die Evaluation erwartet werden. Im nächsten Schritt wählen sie entweder ein für den konkreten Evaluationsgegenstand – die zu untersuchende Maßnahme, das Produkt oder Projekt – geeignetes und bereits vorhandenes logisches Modell bzw. Wirkungsmodell aus oder, falls nicht vorhanden, entwickeln ein logisches Modell (bzw. Wirkungsmodell), das die (vermutete) Wirkungskette des Evaluationsgegenstandes abbildet.

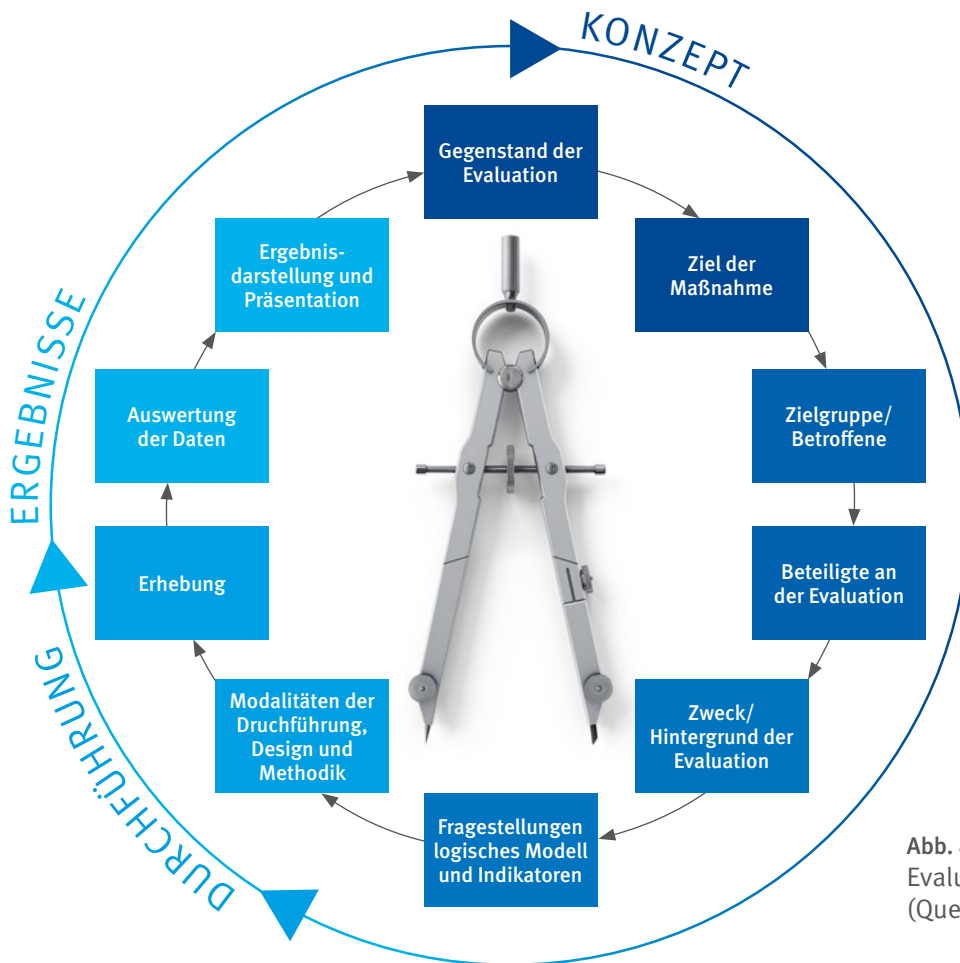


Abb. 4
Evaluationszirkel
(Quelle: eigene Darstellung)

Dieses Modell bildet die logisch aufeinanderfolgenden Schritte ab, anhand derer die zu evaluierende Maßnahme das mit ihr verknüpfte Ziel zu erreichen versucht. Entlang eines solchen logischen bzw. Wirkungsmodells lassen sich dann Indikatoren ableiten, anhand derer sich die schrittweise Zielerreichung untersuchen lässt. Das bedeutet auch, dass Evaluatorinnen und Evaluatoren, die ihre Fragen stellen, die Maßnahmenentwicklung unterstützen können, wenn sie schon so frühzeitig in ein Projekt einbezogen werden. Durch ihr systematisches Vorgehen und ihre zielgerichteten Fragen tragen Evaluatorinnen und Evaluatoren i. d. R. dazu bei, den Fokus und die Ziele von Maßnahmen, Projekten oder Produkten zu schärfen und auf den Punkt zu bringen.

Anhand des folgenden Beispiels eines Seminars zum Thema „Rücken-fit am Arbeitsplatz“ wird einerseits erläutert, anhand welcher Fragen das Evaluationskonzept entwickelt wird und andererseits gezeigt, wie das Evaluationskonzept in der Folge umgesetzt und Ergebnisse abgeleitet werden.

8.1 Den Gegenstand (Maßnahme, Produkt, Projekt) der Evaluation verstehen

In aller Regel geht der Entwicklung einer Maßnahme, eines Produkts oder Projekts eine Bedarfs- und Umfeldanalyse voraus. Die Bedarfs- und Umfeldanalyse ermittelt unter Berücksichtigung bereits vorhandener Angebote, welcher Bedarf für welche Zielgruppe besteht und klärt die Interessen und Erwartungen der verschiedenen Stakeholder an die Maßnahme, das Projekt oder Produkt. Durch die Bedarfs- und Umfeldanalyse können Ressourcen effizient eingesetzt werden, d. h. in der richtigen Zielgruppe und in der richtigen Höhe^[4]. Methoden, die in der Bedarfs- und Umfeldanalyse eingesetzt werden können, sind z. B. Stakeholderanalysen und Gespräche mit den verschiedenen Stakeholdern, eine Doer/Non-Doer Analyse^[29] oder die Erstellung eines Problembaums und Lösungsbaums^[4]. Durch die Bedarfs- und Umfeldanalyse werden wichtige Daten und Informationen gesammelt und generiert, die auch für Evaluatorinnen und Evaluatoren von großem Interesse sind, um den Gegenstand der Evaluation zu verstehen.

Im ersten Schritt des Evaluationszirkels geht es also darum, die Maßnahme, das Produkt oder das Projekt, das evaluiert werden soll, genau zu erfassen. Es ist daher erforderlich, dieses so ausführlich wie möglich zu beschreiben. Mögliche Fragen lauten: Was genau soll evaluiert werden (z. B. Seminar, Medien, Kampagne, Intervention ...)? Aus welchen einzelnen Bestandteilen/Modulen setzt sich die Maßnahme, das Produkt oder das Projekt zusammen? Welche einzelnen Themen/Inhalte werden bearbeitet?

Beispiel

Das Seminar „Rücken-fit am Arbeitsplatz“ setzt sich aus drei Modulen à drei Tagen zusammen. Inhalte sind: theoretische Grundlagen, Kennen und Anwenden von verhältnis- und verhaltenspräventiven Einzelmaßnahmen zur Förderung der individuellen Rückengesundheit.

8.2 Ziele der Maßnahme, des Produkts oder Projekts verstehen

Im zweiten Schritt interessieren die Ziele, die mit der Maßnahme, dem Produkt oder dem Projekt, das evaluiert werden soll, verfolgt werden. Aus welchem Bedarf heraus wurde die Maßnahme, das Produkt oder Projekt entwickelt bzw. welches Problem sollte damit gelöst werden? Welche Ziele werden mit der Maßnahme, dem Produkt oder Projekt verfolgt? Was möchte man erreichen? Welche Teilziele werden mit der Maßnahme, dem Produkt oder Projekt verfolgt? Was sind kurzfristige und was sind strategische, langfristige Ziele? Wann sollen die Ziele erreicht werden?

Beispiel

Muskel-Skelett-Erkrankungen sind eine der Hauptursachen von Arbeitsunfähigkeit und daher für die Unternehmen mit hohen Kosten verbunden. Ziel des Seminars „Rücken-fit am Arbeitsplatz“ ist der Wissensserwerb zum Thema Rückengesundheit sowie das Kennen und Anwenden von Handlungsstrategien zur Prävention von Muskel-Skelett-Beanspruchungen. Insgesamt soll die Gesundheitskompetenz der Teilnehmenden erhöht und das allgemeine Wohlbefinden und die individuelle Rückengesundheit gesteigert werden.

8.3 Zielgruppe/Betroffene der Maßnahme, des Produkts oder Projekts

Im dritten Schritt geht es darum, die Zielgruppe der Maßnahme, des Produkts oder Projekts zu ermitteln. Wer ist die Zielgruppe der Maßnahme, des Produkts oder Projekts (z. B. Beschäftigte, Leitungen, Schülerinnen und Schüler, Sifas, Führungskräfte)? Wie groß ist die Zielgruppe? Wie wird die Maßnahme, das Produkt oder Projekt im Unternehmen/bei der Zielgruppe genau umgesetzt bzw. wie soll sie/es umgesetzt werden? Wie ist der aktuelle Umsetzungsstand? Wer soll oder könnte darüber hinaus von der Maßnahme, dem Produkt oder Projekt erreicht werden (Nebenwirkung)? Auf welchem Weg wird die Zielgruppe der Maßnahme, des Produkts oder Projekts erreicht?

Beispiel

Zielgruppe der Maßnahme „Rücken-fit am Arbeitsplatz“ sind Beschäftigte in Berufen, die durch Rückenbelastungen gekennzeichnet sind. Konkret sind es davon diejenigen, die am Seminar teilnehmen. Die Seminarteilnehmenden sollen verhaltens- und verhältnispräventive Maßnahmen im Betrieb erlernen, anwenden und im besten Fall auch an weitere Kolleginnen und Kollegen weitergeben (Nebenwirkung). Die Zielgruppe wird über die Teilnahme am Seminar erreicht.

8.4 Beteiligte an der Evaluation

Der vierte Schritt hilft zu ermitteln, wer im Rahmen der Evaluation zu beteiligen bzw. zu informieren ist und somit frühzeitig einbezogen werden sollte und mit wem Abstimmungen erforderlich sind. Wer ist der bzw. die Auftraggebende für die Evaluation? Wer sind die Nutzerinnen und Nutzer der Evaluationsergebnisse?

Beispiel

Beteiligte können sein: Der bzw. die Auftraggebende, der bzw. die aufgrund der Evaluationsergebnisse entscheiden möchte, ob das Seminar weitergeführt oder verändert werden soll, die Fachabteilung, die das Seminar entwickelt hat und die Dozentinnen bzw. Dozenten, die das Seminarkonzept umsetzen. Nutzerinnen und Nutzer sind neben dem bzw. der Auftraggebenden z. B. auch die Fachabteilungen, die bei der erneuten Konzeption eines Rückentrainings die Evaluationsergebnisse berücksichtigen können.

An dieser Stelle ist auch zu überlegen, wie viele Personen, die der Zielgruppe der Maßnahme angehören und an der Maßnahme teilgenommen haben, für die Evaluation zur Verfügung stehen. Können alle, die an der Maßnahmendurchführung teilgenommen haben, auch im Rahmen der Evaluation berücksichtigt werden oder muss für die Evaluation eine Stichprobe aus dieser Gruppe gezogen werden? Dies hängt sowohl von den zur Evaluation verwendeten Methoden ab (siehe Kap. 8.7) als auch von den durch die Evaluation zu beantwortenden Fragestellungen und den jeweiligen Rahmenbedingungen der Evaluation.

Beispiel

Im Beispiel des Seminars kann zur Evaluation eine Vollerhebung vorgenommen werden, d. h. in der Evaluation kann auf alle Seminarteilnehmenden und damit die Grundgesamtheit zurückgegriffen werden.

EXKURS

Stichprobe und Repräsentativität

In der empirischen Sozialforschung wird unter Stichprobe eine Teilmenge einer Grundgesamtheit verstanden, die unter bestimmten Gesichtspunkten ausgewählt wurde. Im Kontext der Unfallversicherungsträger kann die Grundgesamtheit beispielsweise aus allen Unternehmen Deutschlands oder aus allen versicherten Betrieben einer bestimmten Branche bestehen. Werden alle Fälle aus der Grundgesamtheit untersucht, spricht man von einer **Vollerhebung oder Totalerhebung**. Eine Vollerhebung liegt auch vor, wenn bei einer Seminarevaluation alle Teilnehmenden des Seminars befragt werden oder wenn eine Abteilung eines Betriebs eine Maßnahme bekommt und alle Mitarbeitenden dieser Abteilung im Anschluss dazu befragt werden. Die an der Maßnahme Teilnehmenden sind jeweils die Grundgesamtheit.

Die Konzentration auf eine Stichprobe ist i. d. R. notwendig, weil in den allermeisten Fällen eine Berücksichtigung der gesamten Grundgesamtheit zu kostspielig und aufwendig wäre. Es wird angenommen,

dass eine **Teilgesamtheit (die Stichprobe)** in den relevanten Merkmalen eine ähnliche Struktur aufweist, wie die Grundgesamtheit und sich daher aus den Ergebnissen der Untersuchung der Teilgesamtheit (Stichprobe) Aussagen über die Grundgesamtheit ableiten lassen. Diese „Repräsentativität“ ist allerdings nur zu erwarten, wenn eine ausreichend große Stichprobe zufällig gezogen wird. Man erwartet dann, dass die Stichprobe eine „verkleinerte Abbildung“ der Grundgesamtheit darstellt und die gleichen Verteilungen und statistischen Beziehungen zwischen Variablen aufweist^[30].

Der Begriff der **Repräsentativität** ist allerdings kein statistischer Fachbegriff und stellt kein klar definiertes Gütekriterium dar^[30]. Die häufig mit diesem Begriff assoziierte Vorstellung, dass eine Stichprobe ein exaktes Abbild der Grundgesamtheit darstellen könnte, ist falsch. „Der Begriff ‚repräsentativ‘ sollte, wenn er für Zufallsstichproben verwendet wird, in dem eingeschränkten Sinn verwendet werden, dass mit diesen

statistisch begründete Schlussfolgerungen auf die Grundgesamtheit möglich sind“ [30, S. 137].

Grundsätzlich stellt sich immer die Frage nach der „**richtigen**“ **Stichprobengröße** – also die Frage nach der Anzahl an Personen, die z. B. befragt werden sollen. Die Antwort auf diese Frage hängt immer auch von der konkreten Fragestellung, der Zielgruppe der Evaluation, den Zielsetzungen der Evaluation sowie von Kosten-Nutzen Abwägungen ab. In manchen Fällen reicht schon die Befragung von wenigen, gut ausgewählten Personen (z. B. Expertinnen bzw. Experten), um relevante und aufschlussreiche Informationen zum Untersuchungsgegenstand zu erfahren.

Übliche Stichprobengrößen bei bevölkerungsweiten Umfragen (z. B. Wahlumfragen oder auch kommerzielle Studien in der Markt- und Meinungsforschung) liegen zwischen 1000 und 3000 Fällen^[30]. Für spezifischere Grundgesamtheiten (z. B. Mitglieder einer Universität oder eines Verbandes) reichen in der Regel Stichprobengrößen zwischen 300 und 1000 Fällen. Sind statistische Analysen geplant, empfiehlt sich eine Mindestgröße von 300 Fällen^[30]. In vielen Fällen sind jedoch schon die Grundgesamtheiten kleiner, sodass die der Evaluation zur Verfügung stehenden Stichproben noch kleiner sind. Zur Berechnung der Mindeststichprobengröße gibt es verschiedene Analyse- und Berechnungsverfahren (z. B. Poweranalysen), die angewandt werden können.

Stichprobenziehung

Ist eine (repräsentative) Stichprobe gewünscht oder gefordert, z. B. um verlässliche Aussagen über die Grundgesamtheit, etwa alle Unternehmen in Deutschland, zu erhalten, gibt es verschiedene Möglichkeiten der Stichprobenziehung. Am Beispiel aller Unternehmen Deutschlands werden im Folgenden einige Verfahren kurz vorgestellt.

Bei einer **Zufallsstichprobe** erfolgt die Auswahl der Unternehmen aus der Grundgesamtheit aller Unternehmen mithilfe von Zufallszahlen oder einem Losverfahren. Eine solche Stichprobe wird als global

repräsentativ bezeichnet, d. h., sie ist repräsentativ bezüglich aller Merkmale.

Bei einer **Klumpenstichprobe** werden mehrere zufällig ausgewählte Teilmengen, z. B. alle Unternehmen aus drei Regionen, vollständig untersucht. Die Generalisierbarkeit der Ergebnisse auf die Grundgesamtheit aller Unternehmen ist in diesem Fall abhängig von Regionsdifferenzen und davon, wie gut die Unternehmen in den ausgewählten Regionen die Gesamtheit der Unternehmen aller Regionen repräsentieren.

Bei einer **geschichteten Stichprobe** wird zunächst eine Schichtung der Grundgesamtheit nach relevanten Merkmalen (z. B. Unternehmensgröße, Standort, Branchenzugehörigkeit) vorgenommen. Anschließend werden Stichprobenumfänge pro Schicht festgelegt und in jeder Schicht eine Zufallsstichprobe im entsprechenden Umfang gezogen. Dadurch werden extreme Stichproben verhindert, die beispielsweise zufällig fast nur Elemente aus einer Schicht enthalten. Eine solche Stichprobe wird als spezifisch repräsentativ bezeichnet, d. h., sie ist repräsentativ bezüglich bestimmter relevanter Merkmale (den Schichten).

Von einer **Ad-hoc-Stichprobe** (auch anfallende Stichprobe) spricht man, wenn keine Stichprobenziehung erfolgt, sondern beispielsweise gut erreichbare Unternehmen untersucht werden oder Unternehmen, die sich für eine Befragung zur Verfügung stellen (Selbstselektion). Diese Stichproben weisen systematische Fehler im Auswahlverfahren auf und sind nicht repräsentativ für die Grundgesamtheit. Die Verlässlichkeit der auf dieser Basis getroffenen Aussagen ist zweifelhaft und Verallgemeinerungen auf die Grundgesamtheit sind nicht möglich bzw. sehr unsicher.

In der Forschungspraxis ist eine geschichtete Zufallsstichprobe häufig ein guter Kompromiss (z. B. nach Unternehmensgröße und/oder Branche). Dieses Vorgehen ist praktikabler als eine reine Zufallsstichprobe und systematische Fehler, wie sie durch reine Ad-hoc-Stichproben entstehen würden, werden eingedämmt.

8.5 Zweck/Hintergrund der Evaluation

Im fünften Schritt interessiert, warum evaluiert werden soll. Denn Evaluationen erfüllen immer einen bestimmten Zweck in Bezug auf die Maßnahme oder im Rahmen eines Projekts. Evaluationen können sich an den Zielen oder an der Durchführung einer Maßnahme orientieren und beantworten eine oder mehrere zuvor festgelegte Fragestellungen. Die Ziele der Maßnahme müssen operationalisierbar und messbar aufgestellt sein. Die Zwecke der Evaluation müssen vor Beginn geklärt werden.

In diesem Schritt interessierende Fragen sind: Warum soll evaluiert werden? Was war der Auslöser für die Entscheidung, eine Evaluation durchzuführen? Was möchte man über die Maßnahme wissen? Was ist der Hintergrund (z. B. politische Vorgabe, Entscheidungsgrundlage, Optimierung)? Wie bzw. wofür sollen die Ergebnisse der Evaluation genutzt werden? Welche Ergebnisse/Wirkungen werden durch die Evaluation erwartet?

Je nachdem, wie die oben genannten Fragen beantwortet werden, verfolgt eine Evaluation hauptsächlich einen oder mehrere der folgenden Zwecke:

- **Erkenntnisgewinn:** Die Evaluation zeigt mithilfe von wissenschaftlichen Methoden auf, ob neu entwickelte oder modifizierte Maßnahmen hinsichtlich der Zielerreichung wirksam sind.
- **Optimierung:** Die Evaluation ermittelt die Stärken und Schwächen der Maßnahme, wie sich diese ausbauen bzw. beseitigen lassen und ob die Maßnahme modifiziert werden sollte. Evaluation in ihrer Optimierungsfunktion setzt die Bereitschaft zur kritischen Reflektion über die Eignung der vorhandenen Maßnahme/Prozesse zur Zielerreichung voraus und trägt damit auch zu kontinuierlichen Lern- und Verbesserungsprozessen und einer Fehlerkultur bei.
- **Bewertung und Entscheidung:** Die Evaluation ermöglicht die Bewertung einer Maßnahme und soll damit die Entscheidung erleichtern, ob eine bestimmte Maßnahme gefördert, weiterentwickelt oder eingestellt werden soll oder ob andere Maßnahmen erforderlich sind, um die gewünschten Ziele zu erreichen. Ist eine Maßnahme erfolgreich, kann auch geprüft werden, ob diese auf andere Bereiche/Branchen übertragbar ist.

- **Kontrolle:** Die Evaluation überwacht die Durchführung einer Maßnahme und stellt fest, ob die Umsetzung korrekt erfolgt und ob sie zu ihrer intendierten Wirkung führt. Evaluation kann dabei z. B. den Transfer von der Theorie in die Praxis überprüfen.
- **Legitimation:** Die Evaluation rechtfertigt den Einsatz von Maßnahmen und entsprechender Ressourcen durch den Nachweis ihrer Effektivität.

Beispiel

In unserem Beispiel soll die Evaluation zum einen Stärken und Schwächen des Seminarkonzepts identifizieren (formative Evaluation). Die Evaluation verfolgt also den Zweck der Optimierung. Zum anderen soll die Wirksamkeit des Seminars ermittelt werden (summative Evaluation). Dies dient dem Zweck der Legitimation. Schließlich soll eine Entscheidung getroffen werden, ob das Seminar weitergeführt wird (Entscheidungsfunktion).

Um die Professionalität und Qualität von Evaluationen zu sichern sowie die Interessen aller Betroffenen und Beteiligten zu wahren, sollen Evaluationen sich an den anerkannten Standards guter Evaluation orientieren, die von der DeGEval formuliert wurden^[13]. Die folgenden vier grundlegenden Eigenschaften (Standards) sollte grundsätzlich jede Evaluation aufweisen:

- **Nützlichkeit:** Die Nützlichkeitsstandards sollen gewährleisten, dass Evaluationen nicht grundlos durchgeführt werden. Evaluation sollte sich immer am Evaluationszweck und dem Informationsbedarf der Nutzerinnen und Nutzer orientieren. Es ist zu klären, warum evaluiert wird und wie die Evaluationsergebnisse genutzt werden sollen. Die Evaluation wird transparent und nachvollziehbar durchgeführt; Evaluatorinnen und Evaluatoren verpflichten sich zu einer transparenten und vollständigen Berichterstattung.
- **Durchführbarkeit:** Jede Evaluation wird realistisch und ökonomisch vorbereitet. Das Vorgehen entspricht einem angemessenen Verhältnis von Aufwand und Nutzen und wird so kommuniziert, dass eine möglichst hohe Akzeptanz aller Beteiligten erreicht wird.

- **Fairness:** Durch einen respektvollen Umgang und die Berücksichtigung individueller Rechte stellt Evaluation Fairness für alle Beteiligten sicher. Im Evaluationsprozess und in den Evaluationsberichten spiegelt sich die unparteiische Position der Evaluatorinnen und Evaluatoren wider (Neutralität und Objektivität). Die Evaluationsergebnisse und -berichte werden allen Beteiligten und Betroffenen soweit wie möglich zugänglich gemacht.
- **Genauigkeit:** Das Messinstrument beziehungsweise die Methode der Evaluation sollte zuverlässige (reliable) und gültige (valide) Informationen liefern. Der Evaluationsgegenstand muss klar definiert und der Kontext (z. B. Betrieb/Organisation), in dem eine Maßnahme durchgeführt wird, hinsichtlich aller bekannter Faktoren, die die Maßnahmendurchführung beeinflussen können, bekannt und analysiert sein. Die im Rahmen der Evaluation genutzten Informationsquellen müssen hinreichend dokumentiert werden. Um die Fragestellung einer Evaluation angemessen beantworten zu können, müssen gewonnene qualitative und quantitative Informationen nach fachlichen Maßstäben analysiert und auf Fehlerquellen geprüft werden.

Der Nutzen der Evaluation von Präventionsleistungen der Unfallversicherungsträger kann darin liegen, dass

- Wirkungen von Präventionsmaßnahmen geprüft und dokumentiert werden,
- auch unbeabsichtigte Wirkungen der Präventionsmaßnahme erkannt und bewertet werden,
- die Selbstverwaltung und die Öffentlichkeit über die Wirksamkeit von Präventionsmaßnahmen informiert werden kann,
- neue oder bestehende Präventionsmaßnahmen verbessert werden können,
- Alternativen zur Präventionsmaßnahme ermittelt werden können,
- durch eine transparente Evaluation ein Lernprozess aller Beteiligten in Gang gesetzt werden kann,
- Argumente zur Begründung von Präventionsmaßnahmen zusammengetragen werden können,

- Hinweise auf notwendige Anpassungen an globale Entwicklungen, Trends oder technische Entwicklungen und
- Informationen zur weiteren strategischen Steuerung der Prävention inklusive der personellen und finanziellen Ressourcen gewonnen werden können.

8.6 Ermittlung von Fragestellungen der Evaluation, logischem Modell und Indikatoren

8.6.1 Fragestellungen der Evaluation

Der sechste Schritt dient der Identifikation des Erkenntnisinteresses und damit der Beantwortung der folgenden Fragen: Was möchte man mithilfe der Evaluation über die Maßnahme, das Produkt oder das Projekt erfahren? Welche konkreten Fragestellungen sollen mit der Evaluation beantwortet werden? Welche Aussagen sollen durch die Evaluation getroffen werden können?

Wenn das Erkenntnisinteresse klar ist, wird ein logisches Modell (Wirkungsmodell) zugrunde gelegt oder ggf. entwickelt, das erläutert, wie bzw. über welche Zwischenschritte die Maßnahme das Ziel erreicht.

8.6.2 Das logische Modell (Wirkungsmodell)

Obwohl sie in der Praxis leider häufig nicht erfolgt, wäre die Erstellung eines logischen Modells schon in der Phase der Maßnahmenplanung sehr hilfreich. Denn Präventionsmaßnahmen werden durchgeführt, weil durch sie eine Veränderung erwartet wird. Durch die Veränderung sollen bestimmte Ziele erreicht und dadurch eine Wirkung erzeugt werden. Die Wirkung der Maßnahme entsteht in der Regel schrittweise über verschiedene Stufen oder Schritte einer Wirkungskette. Und Wirkungsketten lassen sich mit sogenannten logischen Modellen beschreiben. Logische Modelle tragen dazu bei, das Ausmaß einer Veränderung, welches durch die Maßnahme erreicht werden kann, realistisch einzuschätzen.¹⁾

¹⁾ Zur Beschreibung von angenommenen Wirkungsketten wird i. d. R. ein Modell erstellt. Im Kontext der gesetzlichen Unfallversicherung wird dieses Modell oft als „Wirkungsmodell“ bezeichnet. In anderen nationalen Einrichtungen und im internationalen Kontext spricht man in diesem Zusammenhang von „logischen Modellen“. Daher benutzen wir hier auch die Begrifflichkeit des „logischen Modells“.

Logische Modelle

- stellen komplexe Zusammenhänge übersichtlich dar,
- verdeutlichen, mit welchen Mitteln, Maßnahmen, Produkten oder Dienstleistungen über welche Zwischenresultate (Stufen/Schritte/Verkettungen) eine maßnahmeninduzierte Veränderung (Projektziel) erreicht werden soll,
- beschreiben damit die expliziten und impliziten Annahmen über die Wirkungsweise einer Maßnahme und
- bilden den zeitlichen Wirkungsverlauf ab.

Ein logisches Modell ist ein wichtiger Bestandteil im Evaluationskonzept. Es ist zentral für die statistische Überprüfung des Verlaufs und der Wirkungsweise einer Präventionsmaßnahme sowie für das Analysieren und Verstehen von geglückter und missglückter Zielerreichung. Logische Modelle erhöhen die Qualität einer Evaluation, sollten immer grafisch abgebildet werden und selbsterklärend sein. So können sie als Kommunikationsgrundlage dienen.

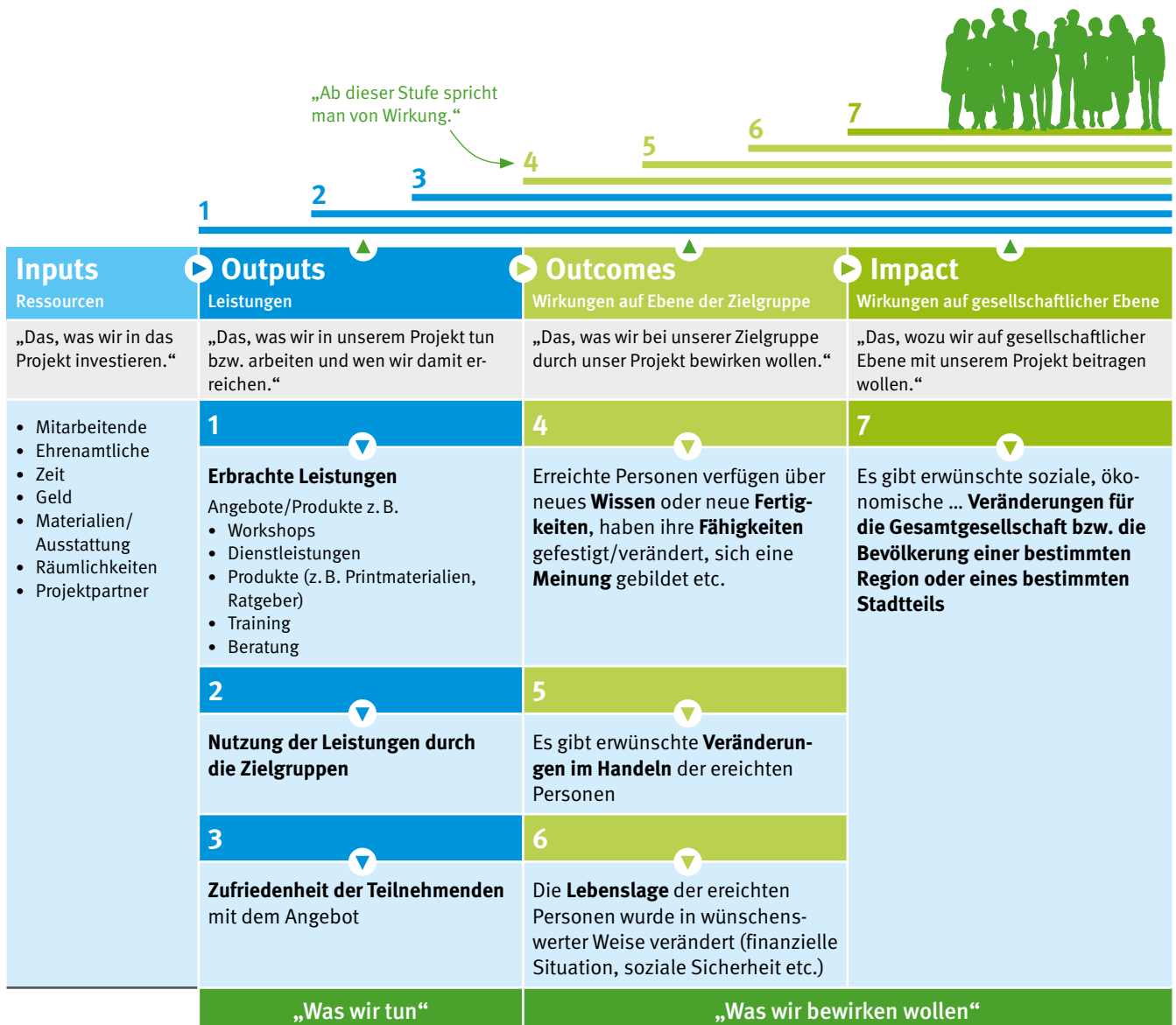


Abb. 5 Logisches Modell (Quelle: Kurz, B., & Kubek, D. (2017). Kursbuch Wirkung. Das Praxishandbuch für alle, die Gutes noch besser tun wollen (4. ed.). Berlin, Gütersloh: Phineo gAG, Bertelsmann Stiftung)

EXKURS

Von Kurz und Kubek^[4] wird ein logisches Modell vorgestellt, das mehrere vorangehende Überlegungen aus der Evaluationsforschung zusammenfasst (siehe Abbildung 5). Es basiert z. B. sehr grundlegend auf dem logischen Modell der W. K. Kellogg Foundation^[32] und greift die Ideen des Vier-Ebenen-Modells von Kirkpatrick^[22] und des Informationsverarbeitungsparadigmas von McGuire^[23] sowie des Qualitätsmodells von Donabedian^[21] und des Prozessorientierten Wirkungsmonitorings^[33] auf.

Die ersten zwei Säulen beschreiben das geplante Vorgehen (Inputs und Outputs) und die letzten zwei Säulen die erwarteten Ergebnisse und Wirkungen (Outcomes und Impacts).

Inputs (= Ressourcen)

Inputs umfassen alle personellen, finanziellen, organisationalen und gesellschaftlichen Ressourcen, die der Maßnahme/dem Programm zur Verfügung stehen und zur Maßnahmen-/Projektumsetzung notwendig sind. Inputs dienen der Erstellung/Durchführung von Leistungen (z. B. Trainings, Technologien, Events, Beratungen, Anreizsysteme), die der Zielgruppe angeboten werden, um die beabsichtigten Veränderungen herbeizuführen.

Outputs

Outputs umfassen die mithilfe der Inputs erstellten direkten Ergebnisse der Maßnahme/der Programmaktivitäten, d. h. die Leistungen für die Zielgruppe sowie die Nutzung der Leistungen durch die Zielgruppe. Die Outputs sind Voraussetzungen für das Erzielen von Wirkungen (Outcomes und Impacts).

Es werden drei Stufen von Outputs unterschieden (Stufen 1–3 der Wirkungslogik, vgl. Abbildung 5):

- Outputs der Stufe 1 sind die Leistungen (Aktivitäten, Produkte, Angebote), die der Zielgruppe geboten werden.
- Outputs der Stufe 2 entsprechen der Nutzung der Leistungen durch die Zielgruppe.
- Die Outputs der Stufe 3 beziehen sich auf die Zufriedenheit der Teilnehmenden aus der Zielgruppe mit den Leistungen. Hintergrund ist die Annahme,

dass die Zufriedenheit der Teilnehmenden mit den Leistungen der Türöffner ist, damit die Teilnehmenden sich in Richtung der intendierten Wirkung entwickeln. Dennoch reicht Zufriedenheit allein noch nicht aus, um tatsächlich und zuverlässig Wirkungen zu erzielen. Es wird aber angenommen, dass die Zufriedenheit mit Leistungen der Maßnahme eine Voraussetzung für eine Veränderung in Richtung der intendierten Wirkungen ist^[4].

Outcomes

Outcomes sind Wirkungen, die durch die Maßnahme, das Projekt oder Produkt erzielt werden sollen und sich auf der Ebene der Zielgruppe vollziehen. Outcomes entstehen durch die Nutzung der Outputs durch die Zielgruppe und ihrer Zufriedenheit damit. Outcomes äußern sich in drei Stufen (Stufen 4 – 6 der Wirkungslogik, vgl. Abbildung 5):

- Outcomes der Stufe 4 zeigen sich z. B. in der Erweiterung von Wissen, Fähigkeiten und Fertigkeiten sowie Einstellungen.
- Outcomes der Stufe 5 äußern sich in Änderungen von Verhalten und Handlungen.
- Outcomes der Stufe 6 manifestieren sich in Veränderungen der Lebenslage der Teilnehmenden an der Maßnahme/dem Projekt in die erwünschte Richtung. Outcomes der Stufe 6 sind z. B. Statusveränderungen oder veränderte betriebliche Zustände. Im logischen Modell der W. K. Kellogg Foundation wird zwischen kurz- (1 bis 3 Jahre) und langfristigen Outcomes (4 bis 6 Jahre) unterschieden^[32].

Impact

Impact beschreibt die beabsichtigte oder unbeabsichtigte Wirkung der Maßnahme/des Programms auf die Zielgruppe, Organisationen, Gemeinden, Systeme, Politik und Gesellschaft. Der Impact ist meist erst nach sieben bis zehn Jahren zu erwarten^[32]. Nicht nur die Maßnahme, das Projekt oder Produkt beeinflussen den Impact, sondern auch viele andere Faktoren haben Einfluss auf gesellschaftliche Veränderungen. Somit ist der kausale Einfluss der Maßnahme, des Projekts oder Produkts auf den Impact kleiner als auf die Outcomes^[4].

Logische Modelle dienen auch dazu, ggf. überzogene Erwartungen von Stakeholdern an die Wirkung von Maßnahmen hinsichtlich ihrer Reichweite in der Wirkungskette zu relativieren. Beispielsweise ist nicht davon auszugehen, dass eine kleine Maßnahme wie z. B. ein Flyer zu UV-Schutz die Neuerkrankungen von Hautkrebs bei Beschäftigten senken wird. Der Flyer allein ist zu „schwach“, um eine Verhaltensänderung herbeizuführen. Achten aber gleichzeitig Führungskräfte konsequent auf sonnensicheres Verhalten ihrer Mitarbeitenden, verhalten sich selber sonnensicher, setzen Belohnungen für sonnensicheres Verhalten ein und informieren über den Flyer hinaus über Folgen von sonnenunsicherem Verhalten, so wächst die Wahrscheinlichkeit, dass die Beschäftigten höhere Stufen des logischen Modells erreichen und ihr Verhalten ändern. Um zu sehen, ob gewünschte Maßnahmenziele mit der geplanten Maßnahme realistisch umsetzbar sind, ist es hilfreich, logische Modelle schon bei der Maßnahmenentwicklung zu erstellen.

Im Vergleich zu Evaluationsmodellen beschränken logische Modelle sich auf konkrete Annahmen über die stufen- oder etappenweise Wirkung einer Maßnahme. Evaluationsmodelle sind i. d. R. umfassender und beziehen oft auch planerische, vorbereitende Maßnahmen sowie die Betrachtung des Kontexts mit ein (siehe z. B. das Modell von Stufflebeam oder das DIEM).

Logische Modelle werden in der wissenschaftlichen Literatur auch als „Theory of Change“ bezeichnet oder, immer dann wenn sie stark an eine bestimmte Maßnahme oder ein Maßnahmenpaket angepasst sind, auch als „Programmlogik“ oder „Programmtheorie“^[31].

8.6.3 Indikatoren

Indikatoren zeigen an, ob die Maßnahme, das Produkt oder Projekt die Zielgruppe erreicht hat und wirkt. D. h., Indikatoren zeigen an, ob Etappenziele, die letztendlich zum erwünschten Ergebnis bei der Zielgruppe führen, erreicht wurden und ob das Projekt auf dem richtigen Weg ist. Durch logische Modelle können schon in der Projektplanung realistische Ober- und Teilziele sowie realistische Indikatoren auf den einzelnen Wirkungsstufen zur Messung des Effekts einer Maßnahme abgeleitet werden. Dazu bietet sich an, das logische Modell unter Beachtung aller Wirkungsstufen von rechts nach links durchzugehen und immer zu fragen „... was getan werden muss beziehungsweise, was passieren muss, um die jeweilige Wirkung bzw. Leistung zu erzielen“^[4]. Die Indikatoren können so im Rahmen der Evaluation operationalisiert und erfasst werden. Wichtig ist, auch zu klären, ob es quantitative oder qualitative Zielvorgaben für die Durchführung bzw. Wirkung der Maßnahme, des Produkts oder Projekts gibt.

Beispiel

Zur Evaluation von Seminaren wird als zugrundeliegendes Modell häufig das Vier-Ebenen-Modell von Kirkpatrick^[22] oder das Informationsverarbeitungsparadigma von McGuire^[23] eingesetzt. Auf Basis des jeweiligen Modells können die Fragestellungen und Indikatoren abgeleitet werden. Im Folgenden werden einige mögliche Indikatoren anhand des Vier-Ebenen-Modells von Kirkpatrick beschrieben:

Reaktionsebene: *Zufriedenheit mit der Maßnahme*

Lernerfolgsebene: *Wissenszuwachs, Akzeptanz und Bereitschaft, die Maßnahme anzuwenden*

Verhaltensebene: *qualitative und quantitative Anwendung der Maßnahme, einschränkende Rahmenbedingungen*

Gesamtresultat: *Wohlbefinden der Teilnehmenden, Häufigkeit der Rückenbeschwerden, Arbeitsbedingungen im Betrieb und betriebliche Kennzahlen (z. B. Fehlzeiten)*

EXKURS

Unfall- und Berufskrankheitenzenzahlen: Eignen sie sich als Indikatoren für die Wirkung von Präventionsmaßnahmen?

Präventionsmaßnahmen können selbstverständlich darauf abzielen, Anzahl und Schwere von Unfällen und Berufskrankheiten zu senken.

Gleichzeitig stellen Unfallzahlen oder Berufskrankheitenzenzahlen jedoch nur schlechte Indikatoren für die Wirkung einer Präventionsmaßnahme dar. Denn hier liegt folgende Annahme zugrunde: Eine Verbesserung der Sicherheit am Arbeitsplatz geht mit einer Senkung der Unfall- und Berufskrankheitenzenzahlen einher. Diese Veränderungen sollen in mittelfristigen Betrachtungen der entsprechenden Statistiken erkennbar werden, die für eine Maßnahme relevant sind. Allerdings gibt es zahlreiche inhaltliche und methodische Gründe, die gegen eine Verwendung dieser Statistiken als Kriterien für die Wirksamkeit von Präventionsmaßnahmen sprechen:

Multiple Einflussfaktoren

Das Problem der Schein-Kausalität ist seit den Ursprüngen der Statistik bekannt. Ein bekanntes Beispiel ist die Annahme, dass der Korrelation zwischen der Zahl der Geburten und der Zahl der Storchpaare in einer Region ein kausaler Zusammenhang zugrunde liegt, obwohl es einen solchen nicht gibt. Die Korrelation zwischen Geburten und Storchpaaren ergibt sich daraus, dass in ländlichen Regionen mehr Störche nisten und dort tendenziell auch mehr Kinder pro Paar geboren werden.

Genauso sind Unfallstatistiken das Ergebnis vieler Faktoren. Verkehrsunfälle werden beispielsweise durch das Wetter, eigenes Verhalten und das anderer Verkehrsteilnehmender sowie die Straßenbeschaffenheit beeinflusst. Weitere Einflussfaktoren auf das Verkehrsunfallgeschehen sind z. B. Schwankungen im Verkehrsaufkommen, die Qualität in der Erfassung von Unfalldaten, rechtliche Vorgaben, konkurrierende Präventionsmaßnahmen anderer Träger der Verkehrssicherheit oder gleichzeitig wirkende Verkehrssicherheitsmaßnahmen von Bund, Ländern, Kommunen. Es ist nicht bekannt, wie alle diese Faktoren zusammenwirken und das Unfallgeschehen im Verkehr beeinflussen.

Der Anteil des Einflusses der Präventionsmaßnahme auf die Unfallzahlen kann also nicht bestimmt und nicht getrennt von den anderen Einflüssen betrachtet werden. Dies ist nur in seltenen kontrollierten und laborartigen Untersuchungen (Experimenten) möglich, wo alle Unfälle erhoben werden, weitere

Einflussfaktoren minimiert oder kontrolliert werden und die Maßnahmen auch im Hinblick auf ihre Evaluation geplant und durchgeführt werden.

Ungenauigkeit von Statistiken

Oft werden Statistiken aus Zahlen einer repräsentativen Stichprobe hochgerechnet. In der Regel beruhen die berichteten Statistiken also nicht auf Vollerhebungen. Daraus können sich Ungenauigkeiten ergeben. Ungenauigkeiten sind bei kleinen Datenmengen (z. B. Fersenbeinbrüche bei Leiterabstürzen) wahrscheinlicher als wenn größere Datenmengen (zum Beispiel alle Arbeitsunfälle) ausgewertet werden.

Zufallsschwankungen

Oft sind Laufzeiten von Maßnahmen zu kurz, um normale Zufallsschwankungen z. B. der Unfallstatistik „auszumitteln“. Jährliche Schwankungen von relativierten Unfallzahlen (ohne Einfluss der Präventionsmaßnahme) liegen bereits bei ein bis vier Prozent. Damit können Veränderungen der Unfallzahlen durch die Maßnahme von diesen Zufallsschwankungen („Rauschen“) überlagert werden.

Zudem gilt: Je kleiner die betrachtete Datenmenge, desto größer der Einfluss von zufälligen Schwankungen. Dieses Problem betrifft auch die Auswertungen tödlicher Unfälle.

Zu geringe Reichweite von Maßnahmen in der Wirkungskette

Oft sind Maßnahmen oder Maßnahmenpakete zu „schwach“, um im logischen Modell Auswirkungen auf der Stufe des Impacts zu erreichen. Insbesondere Maßnahmen, die der Sensibilisierung für ein bestimmtes Thema dienen, erreichen oft nur eine Aufmerksamkeitszuwendung oder die Vermittlung von Wissen. Somit kann eine Wirkung einer Maßnahme oder eines Maßnahmenpakets auf die Senkung von Unfall- und Berufskrankheitenzenzahlen nur in seltensten Fällen und bei sehr starken Maßnahmenpaketen erwartet werden. Bei Berufskrankheitenzenzahlen als Indikator kommt der nächste Punkt erschwerend hinzu.

Lange Latenzzeiten bei Berufskrankheiten

Die Betroffenen sind oftmals jahrelang exponiert, bevor eine Berufskrankheit ausbricht (dies ist beispielsweise bei Muskel-Skelett-Erkrankungen der Fall). Somit können zwischen Exposition und Ausbruch der Krankheit lange Latenzzeiten liegen. Würden Berufskrankheitenzenzahlen zur Bewertung der Wirkung einer Präventionsmaßnahme herangezogen werden, müssten die Evaluationsstudien über entsprechend lange Zeiträume laufen, denn die Maßnahme kann eine Exposition vor der Maßnahme nicht mehr rückgängig machen.

8.7 Modalitäten der Durchführung: Design und Methodik

Im siebten Schritt müssen praktische Fragen zu Möglichkeiten der konkreten Durchführung der Evaluation geklärt werden. Gibt es ein Budget für die Evaluation? Welches Evaluationsdesign bietet sich an? Besteht ggf. die Möglichkeit für Vortests, Erhebungen zu mehreren Messzeitpunkten, Kontrollgruppenbildung? Darüber hinaus muss nun abgeleitet werden, mit welcher Methodik die Zielgruppe erreicht und die Fragestellungen beantwortet werden können.

8.7.1 Evaluationsdesigns

Evaluationsdesigns beschreiben in einem Untersuchungsplan das Vorgehen bei der Datenerhebung und -auswertung. Sie legen fest, wer welche Maßnahme erhält und bei wem wann was gemessen wird. In Abhängigkeit vom Zweck der Evaluation, den zu untersuchenden Zielgruppen, den Fragestellungen der Evaluation und der Verfügbarkeit einer Kontrollgruppe können unterschiedliche Evaluationsdesigns gewählt werden. Vom gewählten Evaluationsdesign hängt ab, welche Schlussfolgerungen aus den Evaluationsergebnissen zur Überprüfung der Wirksamkeit einer Präventionsmaßnahme gezogen werden können. Je nach Art des gewählten Evaluationsdesigns unterscheidet sich die Aussagekraft erheblich.

EXKURS

Evaluationsdesigns

Unterschieden werden Evaluationsdesigns mit Kontrollgruppe von solchen ohne Kontrollgruppe. Ist die Bildung einer Kontrollgruppe nicht per zufälliger Zuordnung möglich, liegt ein quasi-experimentelles Design vor. Wenn die Zuteilung zu den Gruppen zufällig, d. h. randomisiert erfolgt, liegt ein randomisiertes Kontrollgruppendesign vor. Letzteres ermöglicht am ehesten aussagestarke und gültige Schlussfolgerungen^[3].

Die im Bereich der Evaluation von Präventionsmaßnahmen am häufigsten vorkommenden Designs werden im Folgenden kurz vorgestellt:

Post-Messung (One-shot Case Study)

Versuchsgruppe: X – O

Bei diesem Design erfolgt die Befragung bzw. Beobachtung (O), nachdem die Maßnahme (X) durchgeführt wurde. Dabei findet eine einmalige Datenerhebung statt. Mit diesem Evaluationsdesign kann lediglich der Status einer Präventionsmaßnahme zu einem bestimmten Zeitpunkt erfasst werden. Wie der Zustand vor der Maßnahme war und ob weitere Einflüsse an dem Zustandekommen des Ergebnisses beteiligt gewesen sind, kann nicht festgestellt werden.

Prä-Post-Messung mit einer Versuchsgruppe

Versuchsgruppe: O – X – O

Hier wird vor und nach der Maßnahme eine Datenerhebung durchgeführt. Die Prä-Messung wird auch Nullmessung genannt. Auf diese Weise wird erfasst, ob es einen Unterschied zwischen zwei Messzeitpunkten gibt. Damit können jedoch nicht mit Sicherheit weitere Einflüsse benannt oder ausgeschlossen werden. Die eventuell festgestellte Veränderung könnte durch Faktoren zustande gekommen sein, die nicht in der Maßnahme begründet sind.

Prä-Post-Messung mit Versuchs- und Kontrollgruppe

Versuchsgruppe: O – X – O

Kontrollgruppe: O – O

Bei diesem Design erfolgt die Messung vor und nach der Maßnahme sowie im Vergleich zu einer Gruppe, die die Maßnahme nicht erhalten hat. Dabei kann ermittelt bzw. ausgeschlossen werden, ob es andere Einflüsse als die Maßnahme gab, die zu einer Veränderung geführt haben. Idealerweise sollten Personen zufällig, d. h. randomisiert, der Versuchs- bzw. der Kontrollgruppe zugewiesen werden. Dieses Design gilt als „Königsweg“, da es die Probleme und eingeschränkte Aussagekraft der anderen Evaluationsdesigns ausschließt und für die Messung von Veränderungen besonders geeignet ist.

Prä-Post-Messung mit Versuchs- und Warte-Kontrollgruppe

Versuchsgruppe: O – X – O
 Kontrollgruppe: O – O – X – O

Wenn es in der betrieblichen Praxis nicht möglich ist, eine echte Kontrollgruppe zu bilden, d. h. eine Gruppe von Beschäftigten, die die Maßnahme gar nicht erhält, so kann eine sogenannte Warte-Kontrollgruppe gebildet werden. Die Warte-Kontrollgruppe erhält die Maßnahme nachdem die Postmessung in der Versuchsgruppe abgeschlossen ist.

Sollen zwei verschiedene Präventionsmaßnahmen miteinander verglichen werden, entsteht folgendes Design:

Versuchsgruppe 1: O – X1 – O
 Versuchsgruppe 2: O – X2 – O
 Kontrollgruppe: O – O

Follow-up Messungen

Bei all den zuvor dargestellten Designs ist es möglich, die Postmessung um eine oder mehrere, sogenannte Follow-Up-Messungen zu ergänzen, um den Transfer bzw. die Nachhaltigkeit einer Maßnahme zu ermitteln.

Versuchsgruppe: O – X – O – O – ...

8.7.2 Methoden und Instrumente in der Evaluation

Bei der Planung einer Evaluation ist die Frage nach geeigneten Evaluationsmethoden und -instrumenten zentral. Die Methoden beschreiben die systematische Vorgehensweise zur Gewinnung von Erkenntnissen, die Instrumente sind die hierbei eingesetzten Werkzeuge.

In der Regel werden die Methoden im Rahmen des Evaluationskonzepts festgelegt.

Bei der Auswahl der geeigneten Erhebungsmethode/n sollten u. a. folgende Fragen berücksichtigt werden:

- Welcher Zeitraum ist für die Evaluation vorgesehen?
- Gibt es ein Budget für die Evaluation?
- Gibt es bereits Hinweise, ob bestimmte Auswertungen erforderlich sind und wenn ja, welche?
- Mit welchen Methoden oder mit welchem Methodenmix können die Fragestellungen am besten beantwortet werden?
- In welcher Form können die Daten genau erhoben werden? Qualitativ, quantitativ und/oder über Beobachtung?
- Welche Methoden eignen sich am besten für die Zielgruppen (z. B. Papier-/Online-Fragebogen, persönliches Interview) und wie werden die Zielgruppen im Rahmen der Evaluation am besten erreicht (z. B. postalisch, telefonisch, per E-Mail, persönlich)?
- Welche Methoden erzielen ein gutes Verhältnis von Aufwand und Nutzen?

Die empirische Sozialforschung nutzt zahlreiche unterschiedliche Datenerhebungsmethoden. Dabei wird zwischen quantitativen und qualitativen Methoden unterschieden.

Die Informationsschrift der DGUV „Methodenkoffer für die Evaluation“^[34] gibt eine umfassende Übersicht über unterschiedliche quantitative und qualitative Datenerhebungsmethoden in der Evaluationsforschung inklusive möglicher Einsatzgebiete sowie der Vor- und Nachteile der Methoden. Diese Informationsschrift^[34] enthält unter anderem die folgenden Methoden:

- Standardisierter Fragebogen/Test
- Selbst entwickelter Fragebogen
- Standardisiertes/teilstandardisiertes/unstrukturiertes Interview
- Expertengespräche, Experteninterview
- Wissenstest/Lernerfolgskontrolle
- Workshop, Gruppendiskussion, Fokusgruppe, World Café
- Verhaltensbeobachtung, teilnehmende Beobachtung, Begehung
- Messung physiologischer Parameter
- Dokumentenanalyse
- Auswertung von Statistiken/Kennzahlen
- Logfile-Analysen/Webstatistik-Analysen
- Dokumentation

Beispiel

Zur Evaluation von Qualifizierungsmaßnahmen hat sich die Methode der schriftlichen Befragung bewährt. Die Seminarteilnehmerinnen und -teilnehmer erhalten vorher, unmittelbar nach dem Seminar und drei Monate später jeweils einen Fragebogen (Prä-Post-Messung mit Follow-Up ohne Kontrollgruppe). Es werden in diesem Fall überwiegend quantitative Daten erhoben, die alle vier Ebenen des Modells von Kirkpatrick erfassen (siehe Kapitel 8.6.3 Indikatoren).

EXKURS

Quantitative Methoden sind standardisierte Verfahren (z. B. standardisierte Fragebögen oder standardisierte Interviews), die zahlenmäßige Ausprägungen von Merkmalen messen. Deren Auswertung hat das Ziel, auf Basis einer (repräsentativen) Stichprobe verallgemeinerbare Ergebnisse für eine Grundgesamtheit zu generieren und damit Theorien oder Hypothesen zu überprüfen. Um Schlüsse über eine Grundgesamtheit ziehen zu können, ist es notwendig eine möglichst große Anzahl von Personen aus dieser Grundgesamtheit zu befragen. Quantitative Methoden sind im Allgemeinen gut geeignet, wenn es um relativ enge und klar vordefinierte Fragestellungen geht ^[12; 35].

Qualitative Methoden sind vorzuziehen, wenn ein eher offenes, entdeckendes Erkenntnisinteresse besteht. Es werden keine großen Fallzahlen benötigt, da manchmal schon ein einzelner Fall ausreicht, um bestimmte Zusammenhänge aufzuzeigen. Beim qualitativen Vorgehen werden nichtstandardisierte Erhebungsmethoden (z. B. teilstandardisierte oder unstrukturierte Interviews, Experteninterviews oder Gruppendiskussionen) angewendet, in denen überwiegend offene Fragen gestellt werden. Die Befragten können weitgehend frei erzählen. Die Bildung der Stichprobe erfolgt nach theoretischen Gesichtspunkten. Die Stichprobe besteht aus einer kleinen Anzahl von Personen, die für den Forschungsgegenstand relevant sind (z. B. Expertinnen bzw. Experten zu einem bestimmten Thema). Es wird kein Anspruch auf Repräsentativität erhoben. Qualitative Methoden eignen sich besonders für die Untersuchung komplexer oder neuer Fragestellungen, beispielsweise Fragen nach Prozessen und Dynamiken sozialer Systeme, nach Handlungen und Praktiken interessierender Akteurinnen und Akteure oder komplexen sozialen Beziehungen im Kontext von Präventionsmaßnahmen. Insbesondere bei neuen oder komplexen Fragestellungen können qualitative Methoden auch genutzt werden, um Hypothesen zu generieren, die im Anschluss mit Hilfe von quantitativen Methoden überprüft werden ^[12; 35].

Es ist auch möglich und oft sogar sinnvoll, verschiedene Erhebungsmethoden miteinander zu kombinieren oder einzelne Aspekte des Evaluationsgegenstands mit unterschiedlichen Methoden zu untersuchen bzw. unterschiedliche Zielgruppen mit verschiedenen Methoden anzusprechen. In diesem Fall spricht man von einem Methodenmix (engl.: mixed methods).

Die Ergebnisse, die auf Basis verschiedener (z. B. quantitativer und qualitativer) Erhebungsmethoden zum selben Evaluationsgegenstand gewonnen werden, können zusammengebracht und interpretiert werden. Diese Auswertungsstrategie wird als Triangulation bezeichnet. Ein solches Vorgehen versucht, die Vorteile unterschiedlicher Vorgehensweisen zu nutzen.

Die in der Praxis am häufigsten genutzte Methode bei Evaluationen ist die Befragung mittels Fragebogen oder Interview. Es werden zwei Arten von Fragebögen unterschieden: standardisierte Tests bzw. Fragebögen und selbst entwickelte Fragebögen.

Standardisierte Tests oder Fragebögen sind Messinstrumente, die von Wissenschaftlerinnen und Wissenschaftlern für einen bestimmten Zweck entwickelt und getestet wurden, z. B. Intelligenz- oder Persönlichkeitstests. Diese Tests bzw. Fragebögen sind zum Teil frei verfügbar, dürfen „mit Genehmigung“ verwendet werden oder sind käuflich zu erwerben. Sie können z. B. bei der Testzentrale (www.testzentrale.de) oder der Datenbank Assessmentinstrumente des Instituts für Qualitätssicherung in Prävention und Rehabilitation der Deutschen Sporthochschule Köln (www.assessment-info.de) recherchiert werden. Vorteil dieser Verfahren ist, dass sie in der Regel entsprechend der Kriterien Objektivität, Reliabilität und Validität geprüft sind und damit den Gütekriterien entsprechen.

Um Informationen über den Evaluationsgegenstand einzuholen, können **selbst entwickelte Fragebögen** eingesetzt werden. Ein Fragebogen enthält immer eine Abfolge von vorgefassten Fragen oder Aussagen zu einem Themengebiet. Es gibt viele verschiedene Fragearten, die in einem Fragebogen Anwendung finden

können. Generell wird zwischen offenen (freien) und geschlossenen Fragen mit vorgegebenen Antwortkategorien unterschieden. Die Auswahl der Fragearten sowie die Formulierung der Fragen und Antwortmöglichkeiten sollte immer auf das Befragungsziel und die Zielgruppe der Befragung abgestimmt sein. Es ist ratsam, sich bei der Konstruktion Unterstützung von Personen zu holen, die Erfahrung in der Fragebogenentwicklung und der Auswertung von Befragungsdaten haben.

Befragungen mittels Fragebögen können auf unterschiedliche Arten durchgeführt werden, z. B. schriftlich (paper-pencil), online, telefonisch oder persönlich (face-to-face). Die Fragebogenkonstruktion sollte immer auf die geplante Art der Befragung abgestimmt sein.

Ein **Interview** ist eine mündliche Befragung einer Person durch einen oder mehrere Fragenstellende mit dem Ziel, persönliche Informationen und Einstellungen/Meinungen zu ermitteln oder Sachverhalte zu beleuchten. Im Vergleich zur Fragebogenerhebung handelt es sich um eine bewusst hergestellte Gesprächssituation (Kommunikation), bei der es eine Asymmetrie der Frage-Antwort-Zuweisung gibt, d. h., die interviewende Person stellt die Fragen, die interviewte Person beantwortet sie. In der Regel werden Interviews durch geschulte Interviewerinnen und Interviewer geführt, die Erfahrung in der Gesprächsführung haben und sich ihrer Rolle und Verantwortung bewusst sind (z. B. sollte keine eigene Meinung preisgegeben werden, keine Beeinflussung der interviewten Person stattfinden sowie eine freundliche und motivierende Körpersprache eingesetzt werden).

Es werden verschiedene Arten von Interviews unterschieden: standardisierte, teilstandardisierte und unstrukturierte Interviews sowie Experteninterviews. Die verschiedenen Arten von Interviews können wiederum auf unterschiedliche Arten geführt werden, etwa persönlich, telefonisch oder per Video-Konferenz.

In einem **standardisierten Interview** ist die genaue Formulierung der Fragen (einschließlich der Antwortvorgaben) sowie die Reihenfolge der Fragen fest vorgeschrieben. Dadurch werden gleiche Bedingungen für alle Befragten hergestellt, so dass die Daten miteinander verglichen werden können.

Bei **teilstandardisierten Interviews** (auch Leitfadeninterview) kommt ein Interviewleitfaden mit Themenschwerpunkten zum Einsatz, und es werden eher offene Fragen gestellt. Die zentralen Fragen zum Thema sind dabei für alle Befragten gleich, allerdings ist die Frageabfolge veränderbar und vertiefende Fragen sind möglich.

Bei einem **unstrukturierten Interview** existiert kein vorgefertigter Fragebogen oder Leitfaden, sondern ein Stichwortkatalog bzw. ein paar wenige offene Fragen. Die Fragen ergeben sich hauptsächlich aus dem Kontext des Interviews und aus dem Gesprächsverlauf. Nicht alle Befragten erhalten notwendigerweise die gleichen Fragen, so dass man ggf. unterschiedliche, untereinander nicht direkt vergleichbare Interviews erhält.

Eine besondere Art des Interviews ist das **Expertinnen- bzw. Experteninterview**. Die Expertinnen und Experten zeichnen sich durch einen exklusiven Wissensbestand über den Forschungsgegenstand aus. Sie haben privilegierten Zugang zu Informationen, beispielsweise zu Entscheidungsprozessen, oder eine besondere Stellung im Unternehmen (z. B. als Fachkraft für Arbeitssicherheit oder BGM-Koordinatorin oder -Koordinator). Das Expertinnen- bzw. Experteninterview wird in der Regel als Leitfadeninterview mit überwiegend offen formulierten Fragen geführt.

8.8 Datenerhebung

Im achten Schritt wird die Datenerhebung durchgeführt. Für die Wirkungsanalyse und die Prozessevaluation ist es wichtig, Daten nicht nur am Ende eines Projekts zu erheben, sondern auch während der Projektdurchführung. Fragen sind: Wie können Daten erhoben werden, die die Implementierung der Maßnahme, des Produkts oder Projekts dokumentieren und später zur Prozessevaluation herangezogen werden können? Gibt es bereits Informationen, wer bis wann die Ergebnisse der Evaluation benötigt?

Beispiel

Zur besseren Interpretation der Ergebnisse der statistischen Tests wird prozessbegleitend dokumentiert, wie die Präventionsmaßnahmenabwicklung (= die Seminare durchführung) abgelaufen ist, beispielsweise wie viele der angemeldeten Seminarteilnehmerinnen und -nehmer tatsächlich am Seminar teilgenommen haben. Zusätzlich werden Daten zu den Rahmenbedingungen des Seminars erhoben, z. B. ob es Störungen durch äußere Faktoren gegeben hat. Die Ergebnisse der Seminarevaluation sollten bis zur nächsten Jahresplanung vorliegen, so dass Änderungen im Seminar-konzept noch rechtzeitig realisiert werden können.

8.9 Auswertung der Daten

Im neunten Schritt wird die Auswertung der Daten geplant. Bei quantitativen Daten werden mit Hilfe eines Statistikprogramms in der Regel Mittelwerte und Häufigkeiten berechnet. Darüber hinaus lässt sich prüfen, ob bezüglich der Indikatoren eine signifikante Veränderung zwischen den Messzeitpunkten erfolgt ist (statistische Tests auf Mittelwertunterschiede). Qualitative Daten werden inhaltsanalytisch ausgewertet.

Beispiel

Die schriftliche Befragung wird für die drei erhobenen Messzeitpunkte im Hinblick auf die festgelegten Indikatoren (siehe oben) ausgewertet. Im Idealfall sollte ein Wissensanstieg nach dem Seminar erfolgen. Ein erfolgreicher Transfer in den betrieblichen Alltag ist

dann erfolgt, wenn die Anzahl der angewendeten Maßnahmen im betrieblichen Alltag steigt, das Wohlbefinden der Teilnehmerinnen und -nehmer sich verbessert, die Rückenbeschwerden insgesamt weniger werden und sich die Arbeitsbedingungen im Betrieb verbessern (statistische Mittelwertvergleiche über die drei Messzeitpunkte). Sollten die Ergebnisse nicht wie vorhergesagt ausfallen, kann die begleitende Prozessevaluation eventuell eine Antwort geben, z. B. hat der Dozent bzw. die Dozentin tatsächlich zum Thema „Rücken-fit“ referiert oder ist es während der Seminare durchführung zu erheblichen Ablenkungen der Teilnehmerinnen und Teilnehmer von außen gekommen (z. B. Lärm)?

8.10 Ergebnisdarstellung und Präsentation

Die Ergebnisse müssen im zehnten Schritt für alle Beteiligten verständlich dargestellt und rückgemeldet werden, z. B. in Form eines Workshops mit allen Stakeholdern und in einem Evaluationsbericht. Die Fragen, die sich aus dem Erkenntnisinteresse ergeben haben, sollen beantwortet werden. Aus den Ergebnissen sollten gemeinsam mit den Stakeholdern Maßnahmen und Handlungsempfehlungen erarbeitet werden. Der Evaluationsbericht enthält auch Reflektionen darüber, was bei einer nächsten Evaluation besser gemacht werden kann und wie.

Beispiel

Die Ergebnisse der Seminarevaluation werden in einer Power-Point-Präsentation aufbereitet. Diese wird in einem Ergebnisworkshop mit allen Beteiligten (z. B. Fachabteilung, Seminarleitenden) vorgestellt, und es werden Maßnahmen abgeleitet.

9 Herausforderungen und Fallstricke bei Evaluationen

Evaluatorinnen und Evaluatoren begegnen bei ihrer Arbeit zahlreichen Hürden und Herausforderungen. Im Folgenden werden einige Szenarien angerissen, die zu Fallstricken werden können. Es werden Anregungen gegeben, wie sie zu vermeiden sind.

Unzureichende Kenntnis und Erfahrung von Personen, die mit Evaluationen betraut werden

Die Konzeption, Planung und Durchführung von Evaluationen verlangen i. d. R. umfangreiche Kenntnisse und Erfahrungen. Daher ist es unerlässlich Evaluatorinnen und Evaluatoren auszuwählen, die mit dem Ablauf einer Evaluation, Erhebungsmethoden, Auswertungen und Interpretationen von in der Evaluation gewonnenen Daten und Informationen vertraut sind. Dadurch lassen sich häufige Fehler vermeiden, wie z. B.

- die Auswahl von nicht geeigneten Evaluationsmethoden/-instrumenten in Bezug auf die gegebene Fragestellung/Zielgruppe (das ausgewählte Instrument passt nicht zu den gestellten Fragen und kann daher auch keine validen Aussagen liefern)
- nicht eindeutige Fragenformulierung (z. B. zwei Fragen in einer) im Rahmen der Evaluation selbst entwickelten Erhebungsinstrumenten
- unpassende Antwortmöglichkeiten in der Fragebogenentwicklung
- keine oder unzureichende Testung des entwickelten Fragebogens auf Verständlichkeit, Plausibilität und Länge
- keine Baseline-Erhebung, d. h. keine Information über den Ausgangspunkt vor Durchführung der Maßnahme oder des Projekts
- keine Berücksichtigung von Prozessevaluation und Dokumentation der Implementierung einer Maßnahme
- ungeeigneter Zeitpunkt für eine Evaluation (z. B. Evaluation erfolgt zu früh, sodass die Maßnahme noch keine Wirkungen zeigen konnte oder „Kinderkrankheiten“ der Maßnahme noch nicht beseitigt sind).

Dennoch ist der Sachverstand der Evaluationsexpertinnen und -experten alleine nicht ausreichend. Er muss immer auch durch den Sachverstand der jeweiligen Zielgruppe, Branche o. ä. ergänzt werden. Vor allem bei komplexen Evaluationen ist die interdisziplinäre Zusammenarbeit und ein einheitliches Grundverständnis hinsichtlich der geplanten Evaluation für eine erfolgreiche Umsetzung der Evaluation wesentlich.

Fehlende oder unzureichende Bestimmung der Ziele der zu evaluierenden Maßnahme(n)

Es erscheint logisch, dass ohne eine genaue Bestimmung der Ziele der Maßnahme(n) die Zielerreichung nicht überprüft werden kann. Dennoch kann es schwierig sein, im Vorfeld der Maßnahme oder des Projekts exakte und realistische Ziele zu definieren, deren Erreichen dann im Anschluss auch überprüft werden kann. Eine große Herausforderung ergibt sich, wenn sich die Stakeholder zwar über die Wichtigkeit eines Projekts einig sind, aber die Erwartungen hinsichtlich der gewünschten Effekte voneinander abweichen. Häufig haben Auftraggebende auch überzogene Erwartungen an die Effekte einer Maßnahme.

Bei Erwartungen, dass Maßnahmen Effekte am Ende des logischen Modells erreichen (d. h. auf Ebene des Impacts z. B. hinsichtlich betrieblicher Kennzahlen wie der Senkung von Unfallzahlen), kann es sein, dass Ziele nicht messbar sind. Dies ist z. B. der Fall, wenn schon absehbar ist, dass der Effekt der Maßnahme zu schwach sein wird oder die Einflussfaktoren auf den Indikator Unfallzahlen zu zahlreich sind (Multikausalität).

Daher muss der Input, der einer Maßnahmenentwicklung zur Verfügung steht, in Relation zum erwarteten Outcome und Impact stehen. Ist eine Maßnahme „schwach“ und erreicht in ihrer Wirkung „nur“ eine Wissensvermittlung oder dass die Aufmerksamkeit auf einen Sachverhalt gerichtet wird, so führt auch die Evaluation der Maßnahme nicht zu einer Wirkung der Maßnahme auf den höchsten Stufen der Wirkungskette (z. B. Verhaltensänderung (Outcome) oder Veränderung in betrieblichen Kennzahlen (Impact)). Entscheidend ist also, dass eine Maßnahme, die starke Wirkungen haben soll, von vornherein so konzipiert ist, dass sie starke Wirkungen erreichen kann.

Es empfiehlt sich daher, auch die Fragen, die die Evaluation beantworten soll, mit allen Stakeholdern im Vorfeld, d. h. in der Planungsphase, ausreichend abzustimmen und ebenfalls in dieser Phase Klarheit über Bewertungsmaßstäbe zu schaffen (Was bedeuten die Evaluationsergebnisse? Ab wann ist eine Maßnahme „erfolgreich“?). Außerdem ermöglicht eine Stakeholderanalyse im Vorfeld der Evaluation, die Stakeholder vollständig zu identifizieren und deren Erwartungen in Erfahrung zu bringen und ggf. zu relativieren.

Fehlende oder unzureichende Bestimmung des Zwecks der Evaluation

Evaluationen sind keine Garantie dafür, dass die Evaluationsergebnisse weiterverwendet werden und Prozesse anstoßen, die zu einer Verbesserung im Betrieb führen. Damit das „Danach“ nicht wie das „Davor“ ist, muss der Zweck der Evaluation vorher festgelegt werden (Wozu sollen Evaluationsergebnisse genutzt werden?). Evaluation ist kein Selbstzweck. Es soll grundsätzlich nur dann evaluiert werden, wenn sich ein konkreter Nutzen aus der Evaluation ergibt und eine Bereitschaft besteht, aus den gewonnenen Erkenntnissen auch geeignete Maßnahmen oder notwendige Veränderungen abzuleiten und umzusetzen. Andernfalls besteht die Gefahr einer „Alibi-Evaluation“ mit negativen Auswirkungen auf die Akzeptanz künftiger Evaluationen.

Unzureichende Kenntnisse über die oder fehlende Überlegungen zu der/den Zielgruppe(n) der zu evaluierenden Maßnahme(n)

Im Zusammenhang mit der Zielgruppe der Maßnahme sollten Evaluatorinnen und Evaluatoren sich folgende Fragen stellen, die idealerweise bei der Planung der Maßnahme bereits beantwortet sind:

- Welche Zielgruppe(n) soll(en) mit einer Maßnahme erreicht werden?
- Welche Besonderheiten weist die Zielgruppe/weisen die Zielgruppen auf?
- Wie kann die Zielgruppe/können die Zielgruppen angesprochen/erreicht werden?
- Welche Resonanz kann realistisch von der/den Zielgruppe(n) erwartet werden (Über-/Unterforderung vermeiden)?
- Welche Befragungsmethoden sind geeignet? Insbesondere, wenn Mitgliedsunternehmen eines Unfallversicherungsträgers im Rahmen der Evaluation eingebunden werden, müssen Aufwand und Nutzen für die Betriebe wohlüberlegt abgeschätzt werden. Ist es realistisch, dass diese sich ausreichend Zeit für die Evaluation nehmen können? Welchen Mehrwert hat die Teilnahme für sie? Wie können die zu befragenden Personengruppen zur Teilnahme „motiviert“ werden?

Einbindung von Evaluatoreninnen und Evaluatoren, nachdem die Maßnahme schon implementiert wurde

Bei der Planung der Evaluation ist der zeitliche Ablauf der zu evaluierenden Maßnahme entscheidend. Die im Rahmen der Evaluation geplanten Erhebungen müssen zeitlich auf den Ablauf der Maßnahme(n) abgestimmt werden. Hier ist es wichtig, zu den richtigen Zeitpunkten die geeigneten Methoden zur Evaluation anzuwenden. Werden Evaluatoreninnen und Evaluatoren beispielsweise erst eingebunden, wenn die Maßnahmenimplementierung schon vorbei ist, ist die Chance auf eine Prozessevaluation vertan. Damit gehen wichtige Erkenntnisse zur Interpretation der Evaluationsergebnisse verloren.

Damit Evaluatoreninnen und Evaluatoren ein Gespür dafür entwickeln können, wie eine Maßnahme, ein Projekt oder Produkt angenommen wird, welche Hindernisse es gibt oder wie es zu ungeplanten Wirkungen kommen konnte, ist es wichtig, dass Evaluatoreninnen und Evaluatoren vor Ort sein und Einblicke in die Zielgruppen und den Kontext der Maßnahmendurchführung bekommen können.

Keine realistische Ressourcenplanung vorhanden

Neben Ressourcen für die eigentliche Evaluation sollten auch Ressourcen für die Maßnahmenentwicklung und Planung der Maßnahmenimplementierung eingeplant werden. Hier hat sich ein partizipatives, aber auch personalintensives Vorgehen bewährt^[19]. Und auch die Prozessevaluation beansprucht Personalressourcen. Daher setzt eine erfolgreiche Evaluation auch eine realistische Ressourcenplanung (zeitlich, personell, finanziell) für das Evaluationsprojekt voraus.

Keine Baseline-Erhebung vorhanden

Wird die Maßnahme oder das Projekt durchgeführt, ohne dass vorher Baseline-Daten zu den Indikatoren erhoben wurden, erschwert das die Wirkungsanalyse, da keine Vergleichsdaten über die Ausgangssituation vorliegen. I. d. R. werden Baseline-Daten während der Bedarfs- und Umfeldanalyse erhoben. Ist dies versäumt worden oder waren wichtige Indikatoren zu Projektbeginn noch nicht bekannt, können die Teilnehmenden an der Maßnahme/

dem Projekt z. B. gebeten werden auf einer Skala anzugeben, wie sie sich oder ihre Situation hinsichtlich eines Indikators nach der Maßnahmen-/Projektdurchführung einschätzen und wie sie sich hinsichtlich desselben Indikators vor der Maßnahmen-/Projektdurchführung gesehen haben. Eine andere Möglichkeit ist durch bestehende objektive Daten auf die Ausgangssituation zu schließen.

Wirkmechanismen sind nicht bekannt/ein geeignetes logisches Modell bzw. Wirkungsmodell ist nicht vorhanden

Dies erschwert die Formulierung der Fragestellungen, die Identifizierung der Indikatoren und damit auch die Auswahl der geeigneten Evaluationsmethoden. Dass Wirkmechanismen nicht bekannt sind, hängt i. d. R. mit der Komplexität des zu evaluierenden Gegenstands und seines Kontexts zusammen. Findet sich kein passendes logisches Modell bzw. Wirkungsmodell für eine Präventionsmaßnahme in der Forschungsliteratur (siehe auch Kapitel 7), gilt es, eigene Annahmen über die Wirkungsweise einer Präventionsmaßnahme anzustellen, ein logisches Modell zu erarbeiten und in der Projektgruppe zu diskutieren.

Hohe Komplexität des zu evaluierenden Gegenstands und seines Kontexts

Der zu evaluierende Gegenstand, z. B. eine Maßnahme, ein Maßnahmenbündel oder auch ein Projekt, ist komplex. Bei komplexen Systemen sind zahlreiche Systemelemente stark miteinander verknüpft. Sie werden darüber hinaus noch von Geschehnissen der Umwelt beeinflusst, entwickeln sich ständig weiter oder verändern sich.

Vier zentrale Merkmale bzw. Dimensionen komplexer Maßnahmen sind daher in Anlehnung an Dörner^[36] Menge und Vielfalt, Vernetzung, Dynamik sowie Intransparenz. Besondere Herausforderungen in komplexen Kontexten sind i. d. R.:

- die Kontrolle der Einflussfaktoren,
- die Prozessevaluation und
- Wechselwirkungen von mehreren Einzelmaßnahmen.

Bei der Evaluation komplexer Maßnahmen können daher folgende Hinweise hilfreich sein bezüglich:

Auftragsklärung

Es ist wichtig, den Evaluationsauftrag eindeutig zu klären. Eine Stakeholderanalyse hilft dabei, alle relevanten Beteiligten zu identifizieren und kennenzulernen. In Gesprächen sollen die Erwartungen der Auftraggebenden und Stakeholder an die Evaluation und an den Evaluationsgegenstand geklärt werden. Die Evaluatorinnen und Evaluatoren sollten Stakeholder offen auf mögliche nicht zu erwartende oder unerwünschte Ergebnisse hinweisen und so realistische Ergebniserwartungen „verhandeln“. Denn bei sehr komplexen Kontexten mit unklaren Wirkmechanismen und Einflussfaktoren kann auch die Evaluation nur begrenzt Fragestellungen beantworten. Dann muss im Vorfeld geklärt werden, welche Ergebnisse eine Evaluation realistisch erbringen kann.

Eine möglichst genaue Auftragsklärung ist wichtig für jede Evaluation. Hilfreich zur Auftragsklärung ist die Checkliste zur Klärung eines Evaluationsauftrags (s. [28] sowie Kapitel 8).

Planung und Durchführung

Es empfiehlt sich, relevante Einzelkomponenten herauszufiltern, Einzelmaßnahmen zu evaluieren und die Ergebnisse in der Gesamtschau im Hinblick auf Wechselbeziehungen zu analysieren.

In komplexen Evaluationen sollte die Methodenvielfalt durch die Kombination qualitativer und quantitativer Methoden genutzt werden. Anschließend werden die Ergebnisse im Rahmen der Triangulation zusammengeführt (vgl. Kap. 8.7).

Wenn es möglich ist, sollte mit logischen Modellen gearbeitet werden (bereits in der Phase der Konzeption der Maßnahme). Dabei wird nicht versucht, Interventionen und Wirkungen über ein Kontrollgruppendesign miteinander in Verbindung zu bringen. Stattdessen werden Interventionsprozesse über Wirkungsketten/logische Modelle bis zu einem Umsetzungspunkt verfolgt, der erwarten lässt, dass die Maßnahme einen Beitrag zum gewünschten Ergebnis leistet. Zum Nachweis des Einflusses der Intervention ist eine genaue und kontinuierliche Beobachtung und Dokumentation der Maßnahmenimplementierung sowie der Kontextfaktoren wichtig

(Prozessevaluation). Man bewegt sich mit der Nutzung von logischen Modellen/Wirkungsmodellen sehr nahe an der Maßnahmenrealität, kann aber keine Effektmaße wie in klassischen Evaluationsstudien berechnen.

Gerade bei komplexen Maßnahmen hat es sich bewährt, die Evaluation schrittweise zu planen und optionale Methoden vorzudenken. Randomisierte Kontrollgruppendesigns finden bei komplexen Evaluationsgegenständen eher keine Verwendung. Betriebe und Organisationen müssen sich heute viel schneller als früher z. B. an veränderte Marktbedingungen anpassen. Daher ist die Durchführung von Maßnahmen und deren Evaluation in Betrieben und Organisationen heutzutage kein isoliertes Ereignis mehr, sondern findet innerhalb eines dynamischen und komplexen Systems statt. Auch das spricht für einen partizipativen Ansatz bei der Maßnahmenentwicklung und der Planung der Implementierung. Wenn die Maßnahmendurchführung gut in das dynamische und komplexe System eines Betriebs eingeflochten wird, dürfte dies der Bereitschaft zur Teilnahme und damit dem anzustoßenden Veränderungsprozess zuträglich sein^[19] (vgl. Kapitel 7).

Ergebnisse

Sollte es zu unvorhersehbaren Entwicklungen im Evaluationsprozess und damit zu nichterwarteten Ergebnissen kommen, ist es im Sinne eines Beitrags zur Fehlerkultur wertvoll, offen mit den Ergebnissen umzugehen und ihr Zustandekommen zu analysieren. Hilfreich ist, wenn in der Organisation bereits eine Lern- und Fehlerkultur besteht.

10 Fazit

Welche Herausforderungen auf Evaluatorinnen und Evaluatoren zukommen können, welche Voraussetzungen eine erfolgreiche Evaluation benötigt und welche Möglichkeiten und Grenzen bestehen, diese Voraussetzungen zu erfüllen, wurde im vorangehenden Kapitel beschrieben. Wenn Evaluation auch nicht alles leisten kann, was möglicherweise von ihr erwartet bzw. gewünscht wird, so liefert sie doch wertvolle Ergebnisse zu Fragestellungen, die für die gesetzliche Unfallversicherung wichtig sind.

Mithilfe von Evaluationen können

- effektive Präventionsmaßnahmen ermittelt werden, die die Sicherheit und Gesundheit der Beschäftigten in den Betrieben und Einrichtungen verbessern,
- bestehende Präventionsmaßnahmen optimiert oder neue Präventionsmaßnahmen konzipiert werden,
- beabsichtigte und unbeabsichtigte Wirkungen von Präventionsmaßnahmen erfasst und alternative Präventionsmaßnahmen hinsichtlich ihrer Wirkungen miteinander verglichen werden und
- Lernprozesse aller an einer Evaluation Beteiligten angestoßen werden.

Somit dient Evaluation der Gewinnung von Informationen über Präventionsmaßnahmen und ist ein strategisches Mittel zur Steuerung der Prävention in der gesetzlichen Unfallversicherung. Dadurch leisten Evaluationen einen erheblichen Beitrag zur Wirtschaftlichkeit, da sie den Trägern der gesetzlichen Unfallversicherung eine passgenaue Präventionsplanung und einen zielgerechten Einsatz ihrer Ressourcen zur Erfüllung ihres gesetzlichen Auftrags ermöglichen.

11 Anlage

Hilfsmittel zur Durchführung von Evaluationen

Das integrative Baukastensystem für Evaluationen der Allgemeinen Unfallversicherungsanstalt (AUVA)

Im integrativen Baukastensystem werden die verschiedenen Elemente einer Evaluation sowie verschiedene Evaluationsansätze übersichtlich zusammengebracht und dargestellt. In Abhängigkeit von der Fragestellung der Evaluation und ihrem Zweck ist die Berücksichtigung bestimmter dargestellter Elemente angezeigt. Im integrativen Baukastensystem verankert ist die Ex-ante Evaluation

(Analyse von Bedarf, Realisierbarkeit und Erfolgsaussichten), die Konzeptevaluation, die Wirkungslogik über Input, Output, Outcome, Impact, das Vier-Ebenen-Modell von Kirkpatrick^[22], die Möglichkeit der Wirksamkeitsprüfung mit den Elementen Prä-Messung, Post-Messung, Follow-Up, Ex-post Evaluation, das Konzept der Umsetzungsgenauigkeit (Fidelity) und der Prozessevaluation (s. Abbildung 6)^[37; 38]. Das integrative Baukastensystem ist ein Hilfsmittel bei der Planung und Durchführung einer Evaluation.

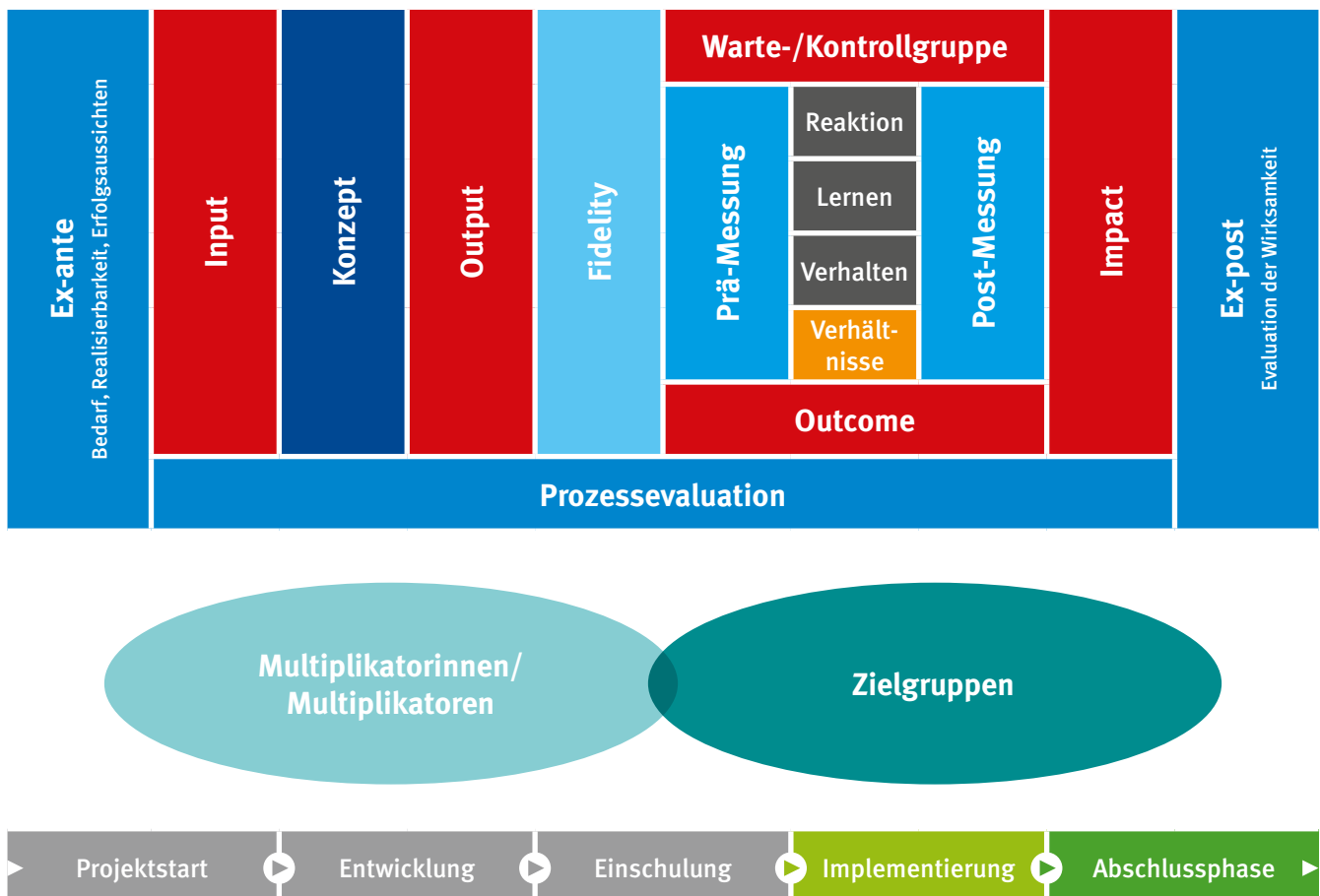


Abb. 6 Integratives Baukastensystem der AUVA (leicht modifizierte Darstellung) (Quelle: Jelenko, M.; Strobach, T.: Evaluation von Präventionsleistungen der Allgemeinen Unfallversicherungsanstalt in Österreich. In: Pfnannstiel, M.; Mehlich, H. (Hrsg.): BGM – Ein Erfolgsfaktor für Unternehmen, S. 819–834. Springer Gabler, Wiesbaden 2018 inhaltlich basierend auf Spiel, G.; Finsterwald, M.; Popper, V.; Hesse, N.: Darstellung des integrativen Baukastensystems für Evaluationen im Präventionsbereich der AUVA. Report Nr. 63. Hrsg.: Allgemeine Unfallversicherungsanstalt, Wien 2013)

Planungshilfe für Evaluationen

Mit dieser tabellarischen Übersicht lässt sich die Evaluation einer Präventionsmaßnahme planen.

Titel der Maßnahme, des Produkts, Projekts, die bzw. das Gegenstand der Evaluation sein soll:	Verantwortlich für die Evaluation:	Zeitraum/-punkt der Evaluation:
---	------------------------------------	---------------------------------

Kurze Beschreibung der zu evaluierenden Maßnahme:

Zweck/Fragestellungen der Evaluation:

Art der Evaluation: Design

Erhebungsmethoden	Erhebungseinheiten und -größe	Indikatoren	Erhebungs-/Messzeitpunkte

Methodische Vorteile:	Methodische Nachteile:
-----------------------	------------------------

Statistisches Auswertungsverfahren

Erläuterungen zu den Begriffen, finden Sie auf der folgenden Seite.

Erläuterung der Begriffe der Planungshilfe für Evaluationen, s. S. 45

Titel der Maßnahme, des Produkts, Projekts, die bzw. das Gegenstand der Evaluation sein soll	Titel der Maßnahme
Verantwortlich für die Evaluation	Institution, Name Zuständige bzw. Zuständiger
Zeitraum/-punkt der Evaluation	
Kurze Beschreibung der zu evaluierenden Maßnahme	Zum Beispiel: An wen richtet sich die Maßnahme? Was ist der Inhalt? Was ist das Ziel?
Zweck/Fragestellung der Evaluation	Was sollte mit der Evaluation in Erfahrung gebracht werden? Welche Daten und Informationen sollen über den Evaluationsgegenstand gewonnen werden?
Art der Evaluation: Design	Zum Beispiel: <ul style="list-style-type: none"> • Post-Messung (One-shot Case Study) • Prä-Post-Messung mit einer Versuchsgruppe • Prä- und Post-Versuchs-Kontrollgruppen-Messung • Prä- und Post-Versuchs-Kontrollgruppen-Messung mit Warte-Kontrollgruppe
Art der Evaluation: Erhebungsmethoden und -instrumente	Zum Beispiel: <ul style="list-style-type: none"> • Standardisierte Tests/Fragebögen • Selbst entwickelte Fragebögen • Standardisiertes Interview • Teilstandardisiertes Interview • Unstrukturiertes Interview
Art der Evaluation: Erhebungseinheiten und -größe	Gemeint sind damit die konkreten Subjekte und/oder Objekte der Erhebungen. Also die konkrete Gruppe der Befragungspersonen im Zusammenhang von z. B. Fragebogenbefragungen oder Interviews (z. B. Erhebungseinheit = FaSis) oder, im Falle von z. B. Dokumentenanalysen, die jeweils ausgewerteten Dokumente (z. B. EE = Unfallakten). Die Erhebungseinheiten ergeben sich also aus der Erhebungsmethode.
Art der Evaluation: Indikatoren	„... empirisch messbare Hilfsgrößen, die nicht direkt erfassbare Phänomene, Zusammenhänge oder Faktoren anzeigen. Sie gestatten Aussagen über den Zustand, die Entwicklung und die Qualitätsmerkmale eines Kriteriums“ Beispiele für Indikatoren: Einkommen, Anzahl der Absolventinnen und Absolventen eines Programms, etc.
Methodische Vorteile	
Methodische Nachteile	
statistisches Auswertungsverfahren	Zum Beispiel: Kreuzkorrelation, Häufigkeitsauszählung, Qualitative Inhaltsanalyse

Projektphasen bei Evaluationsprojekten

Phasen	Tätigkeit
Evaluationsanfrage/ Projektanfrage	Erste Prüfung des Auftrags und Klärung der Auftragsanfrage (mit Checkliste)
	Interne Klärung der Verantwortlichkeiten und Ressourcen für das Projekt (Projektleitung und Beteiligte)
Organisation des Evaluationsprojekts	Erstellung und Abstimmung Evaluationskonzept (inkl. Methodik und Zeitplan und gewünschter Ergebnisdarstellung) als: <ul style="list-style-type: none"> • PowerPoint • ausführliches Konzept (Berichtform in Word)
	Je nach eingesetzter Methodik: Klärung Datenschutz/Einverständnis/Datenarchivierung bzw. -vernichtung
	Festlegung der gewünschten Ergebnisdarstellung mit Auftraggeber <ul style="list-style-type: none"> • Ergebnis-PowerPoint-Präsentation • Kurzveröffentlichung „Management Summary“ • umfassender Abschlussbericht
	Erste Überlegung/Klärung der möglichen Kommunikationsmaßnahmen/Veröffentlichungen <ul style="list-style-type: none"> • Kurzveröffentlichung Print/Online • Fachveröffentlichung Print/Online • Report
	Extern: regelmäßige Kommunikation mit Auftraggeberin bzw. Auftraggeber (Mail, Telefonat, ggf. Projekttreffen) <i>Intern: Kommunikation über Projektstand</i>
Durchführung des Evaluationsprojekts	Planung, Erstellung, Durchführung der abgestimmten Evaluationsmethoden (z. B. Fragebogen, Interview, Workshop etc.), sowie kontinuierliche Datenerfassung/Datensammlung
	Regelmäßiges Projektcontrolling (beinhaltet z. B. Budget, Termin- und Aufgabeneinhaltung) und ggf. Absprache mit Auftraggeberin bzw. Auftraggeber
Abschluss des Evaluationsprojekts	Erstellen der vereinbarten Ergebnisdarstellung für Auftraggeberin bzw. Auftraggeber
	in der Regel kurzes Abschlussgespräch/Feedback mit Auftraggeberin bzw. Auftraggeber
	Umsetzung der geplanten Veröffentlichungen/ggf. Erweiterung der Kommunikationsmaßnahmen (Zustimmung von Auftraggeberin bzw. Auftraggeber einholen)
Dokumentation des Evaluationsprojekts (laufend)	Fortlaufende Erstellung der Projektdokumentation <ul style="list-style-type: none"> • Basisdokumente: Auftrag, Evaluationskonzept • weitere Dokumente während des Evaluationsprozesses, z. B. finale Fragebögen; Interviewleitfaden • Ergebnisdokumentation
	Umgang mit den Daten gemäß Datenschutzvereinbarung



Checkliste zur Klärung eines Evaluationsauftrags bzw. Evaluationsprojekts

Die Beantwortung der nachfolgenden Fragen hilft uns,

- den Evaluationsgegenstand genau kennenzulernen
- Ihr Erkenntnisinteresse zu ermitteln und
- ein Evaluationskonzept zu entwickeln.

1 Maßnahme, Produkt, Projekt, das Gegenstand der Evaluation sein soll

Im ersten Fragenblock geht es darum, die Maßnahme, das Produkt oder das Projekt, das evaluiert werden soll, genau kennenzulernen. Bitte beschreiben Sie dieses daher so ausführlich wie möglich. Gern können Sie uns hierzu auch vorhandene Materialien/Konzeptionen zur Verfügung stellen.

- Was genau soll evaluiert werden? (z. B. Seminar, Medien, Kampagne, Intervention...)
- Aus welchen einzelnen Bestandteilen/Modulen setzt sich die Maßnahme, das Produkt oder das Projekt zusammen? Welche einzelnen Themen/Inhalte werden bearbeitet?

2 Ziele der Maßnahme, des Produkts oder Projekts

Im zweiten Fragenblock interessieren uns die Ziele, die Sie mit der Maßnahme, dem Produkt oder dem Projekt, das Sie evaluieren möchten, verfolgen.

- Aus welchem Bedarf heraus wurde die Maßnahme, das Produkt oder Projekt entwickelt bzw. welches Problem sollte damit gelöst werden?
- Welche Ziele verfolgen Sie mit der Maßnahme, dem Produkt oder Projekt? Was möchten Sie erreichen?
- Welche Teilziele verfolgen Sie mit der Maßnahme, dem Produkt oder Projekt? Was sind kurzfristige und was sind strategische, langfristige Ziele?
- Wann sollen die Ziele erreicht werden?

3 Zielgruppe/Betroffene der Maßnahme, des Produkts oder Projekts

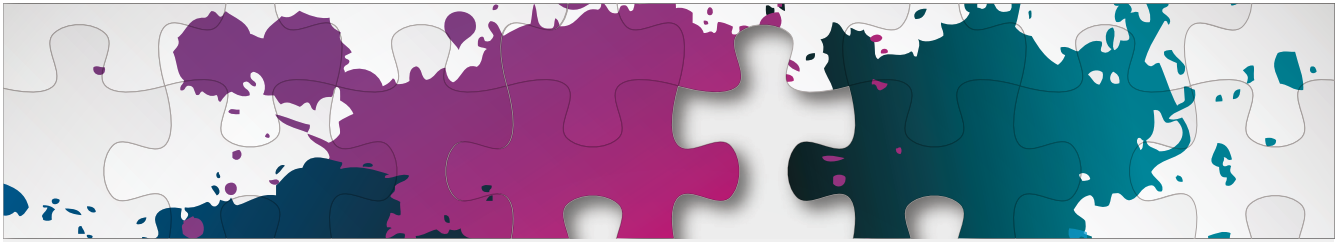
Im dritten Fragenblock geht es darum, die Zielgruppe Ihrer Maßnahme, Ihres Produkts oder Projekts zu ermitteln.

- Wer ist die Zielgruppe der Maßnahme, des Produkts oder Projekts? Wer soll erreicht werden (z. B. Beschäftigte/Leitungen/Schülerinnen bzw. Schüler/Sifas/Führungskräfte...)? Bitte beschreiben Sie die Zielgruppe ggf. näher (z. B. Alter, Geschlecht, Branche, Betriebsgröße, Führungsverantwortung, Besonderheiten, ...)
- Wie groß ist die Zielgruppe?
- Wie wird die Maßnahme, das Produkt oder Projekt im Unternehmen/bei der Zielgruppe genau umgesetzt bzw. wie soll sie/es umgesetzt werden? Wie ist der aktuelle Umsetzungsstand?
- Wer soll oder könnte darüber hinaus von der Maßnahme, dem Produkt oder Projekt erreicht werden (Nebenwirkung)?
- Auf welchem Weg wird die Zielgruppe der Maßnahme, des Produkts oder Projekts erreicht?

4 Beteiligte an der Evaluation

Der vierte Fragenblock hilft uns zu ermitteln, wer im Rahmen der Evaluation zu beteiligen bzw. zu informieren ist und somit frühzeitig einbezogen werden sollte und mit wem Abstimmungen erforderlich sind.

- Wer ist an der Evaluation beteiligt bzw. wer sollte beteiligt oder einbezogen werden (z. B. Betriebsrat etc.)?
- Wer ist die bzw. der Auftraggebende für die Evaluation?
- Wer sind die Nutzerinnen bzw. Nutzer der Evaluationsergebnisse?



5 Zweck/Hintergrund der Evaluation

Im fünften Fragenblock interessiert uns, warum evaluiert werden soll.

- Was war der Auslöser für die Entscheidung, eine Evaluation durchzuführen? Was ist der Hintergrund (z. B. politische Vorgabe, Entscheidungsgrundlage, Optimierung, ...)?
- Wie bzw. wofür sollen die Ergebnisse der Evaluation genutzt werden?

6 Ermittlung von Fragestellungen, logischem Modell und Indikatoren

Im sechsten Fragenblock interessiert uns, was Sie mithilfe der Evaluation über die Maßnahme, Ihr Produkt oder Projekt erfahren möchten, d. h. welches Erkenntnisinteresse Sie an der Evaluation haben. Bitte listen Sie so detailliert wie möglich auf.

Fragestellungen:

- Was möchten Sie mithilfe der Evaluation über Ihre Maßnahme, Ihr Produkt oder Ihr Projekt erfahren?
- Welche konkreten Fragen interessieren Sie, die mit der Evaluation beantwortet werden sollen?
- Welche Aussagen sollen durch die Evaluation getroffen werden können?

Logisches Modell (Wirkungsmodell):

- Welche Annahmen haben Sie hinsichtlich der Schritte/Stufen über die die Maßnahme, das Produkt oder Projekt ihr Ziel erreicht?

Indikatoren:

- Woran würden Sie erkennen, dass die Maßnahme, das Produkt oder Projekt die Zielgruppe erreicht hat und wirkt, d. h. zum erwünschten Ergebnis bei der Zielgruppe geführt hat?
- Gibt es quantitative oder qualitative Zielvorgaben für die Durchführung bzw. Wirkung der Maßnahme, des Produkts oder Projekts?

7 Modalitäten der Durchführung, Design und Methodik

Im siebten Fragenblock interessieren uns, falls zum jetzigen Zeitpunkt schon beantwortbar, einige praktische Fragen zu Möglichkeiten der Durchführung der Evaluation.

- Welcher Zeitraum ist für die Evaluation vorgesehen?
- Welche Methodik oder welcher Methodenmix eignet sich möglicherweise am besten, um die Zielgruppe zu erreichen (z. B. Papier-/Onlinefragebogen, persönliches Interview)?
- Besteht ggf. die Möglichkeit für Vortests, Erhebungen zu mehreren Messzeitpunkten, Kontrollgruppenbildung?
- Wie wird die Zielgruppe im Rahmen der Evaluation am besten erreicht (z. B. postalisch, E-Mail, persönlich, ...)?
- Gibt es ein Budget für die Evaluation?

8 Datenerhebung, Auswertung und Ergebnisdarstellung

Im achten Fragenblock interessieren uns, falls zum jetzigen Zeitpunkt schon beantwortbar, einige Fragen zur Auswertung und Ergebnisdarstellung der Evaluation.

- Gibt es bereits Hinweise, ob bestimmte Auswertungen erforderlich sind und wenn ja, welche?
- Gibt es bereits Informationen, wer bis wann die Ergebnisse der Evaluation benötigt?
- Welche Vorarbeiten (Analysen, Studien, Auswertungen, Materialien, frühere Evaluationen) gibt es ggf. bereits zu Ihrer Maßnahme, Ihrem Produkt oder Projekt?
- Welche Erfahrungen oder Daten zu ähnlichen Maßnahmen, Produkten oder Projekten, gibt es ggf., die für das aktuelle Evaluationsvorhaben nutzbar sind?
- Welche weiteren Informationsquellen sind ggf. vorhanden und nutzbar? (z. B. Statistiken, Veröffentlichungen)

Vielen Dank für Ihre Mitarbeit!

Verzeichnisse

Literatur

- [1] Einführung eines Siebten Buches Sozialgesetzbuch. Bundesgesetzblatt I 43 (1996), S. 1254–1309, http://www.gesetze-im-internet.de/bundesrecht/sgb_7/gesamt.pdf
- [2] Westermann, R.: Merkmale und Varianten von Evaluationen: Überblick und Klassifikation. Z. Psychol. 210 (2002) Nr. 1, S. 4–26
- [3] Leitfaden „In drei Schritten zur Evaluation von Verkehrsprojekten“. Hrsg.: Umweltbundesamt; Gesamtverband der Deutschen Versicherungswirtschaft e. V., Dessau-Roßlau; Berlin 2018
- [4] Kurz, B.; Kubek, D.: Kursbuch Wirkung. Das Praxishandbuch für alle, die Gutes noch besser tun wollen 4. Aufl. Phineo gAG, Bertelsmann Stiftung, Berlin, Gütersloh 2017, <https://www.phineo.org/fuerorganisationen/kursbuch-wirkung/>
- [5] Weber, J.: Controlling. Hrsg.: Gabler Wirtschaftslexikon 2018 <https://wirtschaftslexikon.gabler.de/definition/controlling-30235/version-206821> 20.02.2018 (abgerufen am 07.11.2018)
- [6] Kaplan, R. S.; Norton, D. P.: Balanced scorecard: Strategien erfolgreich umsetzen Schäffer-Poeschel, Verlag für Wirtschaft · Steuern · Recht GmbH, Stuttgart 1997
- [7] Wübbenhorst, K.: Benchmarking. Hrsg.: Gabler Wirtschaftslexikon 2010, <https://wirtschaftslexikon.gabler.de/definition/benchmarking-29988>, 29.07.2010 (abgerufen am 07.11.2018)
- [8] Abschlussbericht. Benchmarking-Projekt „Wirksame Qualifizierung mit wirtschaftlichen Verfahren“. Hrsg.: Deutsche Gesetzliche Unfallversicherung e. V. (DGUV) 2018
- [9] Voigt, K.-I.: Qualitätssicherung. Hrsg.: Gabler Wirtschaftslexikon 2018, <https://wirtschaftslexikon.gabler.de/definition/qualitaetssicherung-44396>, 19.02.2018 (abgerufen am 07.11.2018)
- [10] Beeck, V.: Auditing. Hrsg.: Gabler Wirtschaftslexikon 2018, <https://wirtschaftslexikon.gabler.de/definition/auditing-30394/version-253977>, 15.02.2018 (abgerufen am 07.11.2018)
- [11] Präventionsleistungen der Unfallversicherungsträger der Deutschen Gesetzlichen Unfallversicherung. Stand Dezember 2016 – aktualisiert Mai 2019. Hrsg.: Deutsche Gesetzliche Unfallversicherung e. V. (DGUV), Berlin 2016 <https://publikationen.dguv.de/dguv/pdf/10002/12471.pdf> (abgerufen am 07.11.2018)
- [12] Bortz, J.; Döring, N.: Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften Springer-Verlag, Berlin Heidelberg 2016
- [13] Standards für Evaluation. Hrsg.: Deutsche Gesellschaft für Evaluation (DeGEval), Mainz 2017, https://www.degeval.org/fileadmin/Publikationen/DeGEval_Standards_fuer_Evaluation_-_Erste_Revision__2016_.pdf, 21.09.2016 (abgerufen am 07.11.2018)
- [14] Century, J.; Rudnick, M.; Freeman, C.: A Framework for Measuring Fidelity of Implementation: A Foundation for Shared Language and Accumulation of Knowledge. American Journal of Evaluation 31 (2010) Nr. 2, S. 199–218, <https://doi.org/10.1177/1098214010366173> (abgerufen am 09.05.2019)
- [15] Nielsen, K.; Abildgaard, J. S.: Organizational interventions: A research-based framework for the evaluation of both process and effects. Work Stress 27 (2013) Nr. 3, S. 278–297
- [16] Nielsen, K.; Randall, R.: Opening the black box: Presenting a model for evaluating organizational-level interventions. Eur. J. Work Org. Psychol. 22 (2013) Nr. 5, S. 601–617
- [17] Schmidt, U.: In: 1. Fachgespräch Evaluation – „Standards setzen?!“ – Die Beiträge des Fachgesprächs, gehalten am 25. und 26. März 2010 in Dresden – Deutsche Gesetzliche Unfallversicherung e. V. (DGUV), Berlin, 2010, <http://publikationen.dguv.de/dguv/pdf/10002/report2010-01.pdf>

- [18] Stufflebeam, D. L., The CIPP model for evaluation, presented at the 2003 Annual Conference of the Oregon Program Evaluators Network (OPEN), Portland, Oregon, 2003, <https://de.scribd.com/document/58435354/The-Cipp-Model-for-Evaluation-by-Daniel-L-Stufflebeam>
- [19] von Thiele Schwarz, U.; Lundmark, R.; Hasson, H.: The Dynamic Integrated Evaluation Model (DIEM): Achieving Sustainability in Organizational Intervention through a Participatory Evaluation Approach. *Stress and Health* 32 (2016) Nr. 4, S. 285–293, <https://www.ncbi.nlm.nih.gov/pubmed/27501278>
- [20] Hamacher, W.; Watzl, F.; Eigenstetter, M.; Schwärzel, L.; Keil, U.; Löcher, T.: Evaluation oder „Tun wir die richtigen Dinge? – Tun wir die Dinge richtig?“. Hrsg.: Fachverband Psychologie der Arbeitssicherheit und Gesundheit e. V., Köln 2017, http://www.fv-pasig.de/fileadmin/user_upload/Exp_Evaluation/Reader_PASiG_Evaluation_2017_End.pdf (abgerufen am 23.04.2018)
- [21] Donabedian, A.: Evaluating the Quality of Medical Care. *Milbank Mem. Fund Q. Health Soc.* 44 (1966), S. 166–203
- [22] Kirkpatrick, D.: How to start an objective evaluation of your training program. *Journal of the American Society of Training Directors* 10 (1956), S. 18–22
- [23] McGuire, W. J.: An information processing model of advertising effectiveness. In: *Behavioral and Management Science in Marketing*. (Eds.: Davis, H. L.; Silk, A. J.), Ronald Press, New York, 1978, pp. 156–180
- [24] Lehmann, H.; Töppich, J.: Gesundheitliche Aufklärung als Risikokommunikation. *Erfahrungsbericht aus der AIDS-Prävention. Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz* 52 (2009) Nr. 12, S. 1147–1150
- [25] Pott, E.: AIDS-Prävention in Deutschland. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz* 50 (2007) Nr. 4, S. 422–431
- [26] Wetzstein, A.; Taskan-Karamürsel, E.: Evaluation als Steuerungsinstrument von Kampagnen – Am Beispiel der Präventionskampagne „Risiko raus!“. In: Sachse, Pierre; Ulich, Eberhard (Hrsg.): *Psychologie menschlichen Handelns: Wissen & Denken – Wollen & Tun*, S. 313–327. Pabst Science Publishers, Lengerich 2014
- [27] Hessenmöller, A.; Rahnfeld, M.; Renner, S.; Wetzstein, A.: IAG Report 5/2017. Die Evaluation der Präventionskampagne „Denk an mich. Dein Rücken“ Planung, Durchführung und zentrale Ergebnisse. Hrsg.: Deutsche Gesetzliche Unfallversicherung e. V. (DGUV), Berlin 2017, https://publikationen.dguv.de/dguv/udt_dguv_main.aspx?FDOCUID=26875 (abgerufen am 21.11.2018)
- [28] Checkliste zur Klärung eines Evaluationsauftrags bzw. Evaluationsprojekts. Hrsg.: Institut für Arbeit und Gesundheit der Deutschen Gesetzlichen Unfallversicherung (IAG); Fachbereich Organisation des Arbeitsschutzes der DGUV (FB ORG), Berlin 2018, https://www.dguv.de/medien/fb_org/dokumente/checkliste.pdf, 08.11.2018
- [29] Contzen, N.; Mosler, H. J.: Doer/non-doeer analysis to specify the critical behavioral factors. *Methodological factsheet 5*. Hrsg.: eawag aquatic research 2015, https://docs.wixstatic.com/ugd/accbe3_db150cee6a8349fb8681adcc8e888bd95.pdf
- [30] Diaz-Bone, R.: *Statistik für Soziologen* 2. Aufl. UTB Basics; UVK Verlagsgesellschaft, Konstanz 2013
- [31] Wyatt Knowlton, L.; Phillips, C. C.: *The logic model guidebook. Better strategies for great results*. 2nd. Aufl. SAGE Publications, Inc., Los Angeles, London, New Delhi, Singapore, Washington DC 2013
- [32] *Logic Model Development Guide*. Hrsg.: W.K. Kellogg Foundation, Battle Creek, Michigan 2004, <https://www.wkcf.org/resource-directory/resource/2006/02/wk-kellogg-foundation-logic-model-development-guide> (abgerufen am 27.09.2018)

- [33] Hummelbrunner, R.: Process Monitoring of Impacts. Proposal for a new approach to monitor the implementation of 'Territorial Cooperation' programmes. Hrsg.: INTERACT Point MTEC, Vienna 2006
- [34] DGUV Information (in Erarbeitung): Methodenkoffer für Analysen und Evaluationen. Eine umfangreiche Sammlung von Erhebungsmethoden in der gesetzlichen Unfallversicherung. Hrsg.: Deutsche Gesetzliche Unfallversicherung e. V. (DGUV), Berlin
- [35] Krajic, K.; Reiter, C.; Wimmer, E.; Flecker, J.: Qualitative Erhebungs- und Auswertungsmethoden zur Umsetzung qualitativ hochwertiger Evaluationen – Beitrag zum Baukastensystem für Evaluationen im Präventionsbereich der AUVA. Hrsg.: Universität Wien, Wien 2016
- [36] Dörner, D.: Die Logik des Mißlingens. Strategisches Denken in komplexen Situationen. Rowohlt Verlag GmbH, Reinbek bei Hamburg 2003
- [37] Spiel, G.; Finsterwald, M.; Popper, V.; Hesse, N.: Darstellung des integrativen Baukastensystems für Evaluationen im Präventionsbereich der AUVA. Report Nr. 63. Hrsg.: Allgemeine Unfallversicherungsanstalt, Wien 2013, <https://www.auva.at/cdscontent/load?contentid=10008.544753&version=1547813658>, (abgerufen am 30.04.2018)
- [38] Jelenko, M.; Strobach, T.: Evaluation von Präventionsleistungen der Allgemeinen Unfallversicherungsanstalt in Österreich. In: Pfannstiel, M.; Mehlich, H. (Hrsg.): BGM – Ein Erfolgsfaktor für Unternehmen, S. 819–834. Springer Gabler, Wiesbaden 2018

Weiterführende Literatur der DGUV

- IAG Report 1/2019 Evaluation der RAG-Kampagne „SICHERHEIT! Denk daran, bevor du loslegst.“
Ein Pilotprojekt der BG RCI für die Präventionsstrategie VISION ZERO
 - DGUV Information 206-022 – Verfahren und Methoden im Präventionsfeld „Gesundheit im Betrieb“ – Empfehlungen für Präventionsfachleute, 2019
 - IAG Report 5/2017 Die Evaluation der Präventionskampagne „Denk an mich. Dein Rücken“ Planung, Durchführung und zentrale Ergebnisse
 - IAG Report 3/2017 Evaluation in der gesetzlichen Unfallversicherung – Beispiele aus der Projektarbeit Die Beiträge der Veranstaltungen „Fachgespräch Evaluation“ in Dresden, gehalten in den Jahren 2012, 2014 und 2016
 - IAG Report 1/2017 Erfolgsfaktoren für Arbeitsschutzunterweisungen – Identifikation und Messung
 - DGUV Report 1/2017 Evaluation der DGUV Vorschrift 2, Anlage 2 (Abschlussbericht) Stand der Umsetzung, Anwendbarkeit und Praktikabilität sowie Auswirkungen im Betrieb
 - IAG Report 1/2016 Prävention gegen sexuellen Missbrauch bei Kindern Evaluationsergebnisse des Präventionsprogramms „Gemeinsam · Mädchen und Jungen stärken“ des Kinderschutzbundes Münster
 - IAG-Report 1. Fachgespräch Evaluation – „Standards setzen?!“ – Die Beiträge des Fachgesprächs, gehalten am 25. und 26. März 2010 in Dresden –
- Alle zu finden unter: <https://publikationen.dguv.de/>

Abbildungs- und Tabellenverzeichnis

Abb. 1	Präventionsleistungen der Unfallversicherungsträger	10
Abb. 2	Das Dynamische, Integrierte Evaluationsmodell	18
Abb. 3	Ebenenmodell der Evaluation von Präventionskampagnen	22
Abb. 4	Evaluationszirkel	24
Abb. 5	Logisches Modell	30
Abb. 6	Integratives Baukastensystem der AUVA (leicht modifizierte Darstellung)	44
Tabelle 1:	Präventionsleistungen und mögliche Evaluationsgegenstände.....	11

**Deutsche Gesetzliche
Unfallversicherung e.V. (DGUV)**

Glinkastraße 40
10117 Berlin
Telefon: 030 13001-0 (Zentrale)
Fax: 030 13001-9876
E-Mail: info@dguv.de
Internet: www.dguv.de