

Expertenkreis

Evaluation, Wirksamkeit und Nachhaltigkeit

Fachverband Psychologie für Arbeitssicherheit und
Gesundheit e.V.

Evaluation oder „Tun wir die richtigen Dinge? – Tun wir die Dinge richtig?“

Autoren: Werner Hamacher, Frank Watzl, Monika
Eigenstetter, Lotte Schwärzel, Udo Keil, Tobias Löcher
(Köln, 2017)

Inhaltsverzeichnis

1	Evaluation oder „Tun wir die richtigen Dinge? Tun wir die Dinge richtig?“	2
1.1	Evaluation im Arbeitsschutz verstehbar machen.....	2
1.2	Das CIPP-Modell.....	4
1.2.1	Kontextevaluation (C): „Was muss getan werden?“	4
1.2.2	Inputevaluation (I): „Wie muss es getan werden?“.....	4
1.2.3	Prozessevaluation (P): „Wurde das Programm/ das Verfahren wie vorgesehen umgesetzt?“	5
1.2.4	Produktevaluation (P): „Führt das Programm zum Erfolg?“	5
1.2.5	Fazit	6
2	Evaluation? Wo steht das im Gesetz?.....	6
2.1.1	Arbeitsschutzgesetz: Sicherheit und Gesundheit als Managementprogramm.....	6
2.1.2	Begleitende, wissensbasierte Optimierung: Programmevaluation	7
2.1.3	Kontextevaluation: Was soll erreicht werden?	8
2.1.4	Inputevaluation: Wie soll es erreicht werden?.....	8
2.1.5	Prozessevaluation: Wird es getan?	8
2.1.6	Produktevaluation: Hat es gewirkt?.....	8
2.1.7	Fazit	9
3	Stakeholder-Analysen als Voraussetzung für gelingende Evaluationen.....	10
3.1	Einführung.....	10
3.2	Stakeholder – Definition.....	10
3.3	Einflussmöglichkeiten verschiedener Anspruchsgruppen in Arbeitssicherheit und Gesundheitsschutz 11	
3.4	Fazit	12
4	Prozessevaluation eines Projektes zum betrieblichen Gesundheitsmanagement.....	15
4.1	Gesundheits-Balanced-Scorecard.....	15
4.2	Prozessevaluation	16
4.3	Methodik.....	17
4.4	Ergebnisse.....	17
4.5	Handlungsempfehlungen	18
5	Produktevaluation am Beispiel „Gesundheits-Balanced-Scorecard“	19
5.1	Gesundheits-Balanced-Scorecard.....	19
5.2	Produktevaluation.....	19
5.3	Ergebnisse.....	20
5.4	Diskussion	21
5.5	Zusammenfassung	22
6	Literatur	23
7	Autor*innen	26

1 Evaluation oder „Tun wir die richtigen Dinge? Tun wir die Dinge richtig?“

1.1 Evaluation im Arbeitsschutz verstehbar machen

In der betrieblichen Praxis von Arbeitssicherheit und Gesundheitsschutz spielt Evaluation bislang kaum eine Rolle. Doch erfordert das Arbeitsschutzgesetz ein systematisches Vorgehen und einen ständigen Verbesserungsprozess um Sicherheit und Gesundheit der Beschäftigten zu gewährleisten. Evaluation sollte daher auch ein selbstverständlicher Teil des betrieblichen Arbeitsschutzhandelns werden.

Damit Evaluation selbstverständlicher Teil des betrieblichen Arbeitsschutzhandelns werden kann, muss Evaluation für betriebliche Praktiker wie Unternehmer/ innen, Führungskräfte, Fachkräfte für Arbeitssicherheit, Betriebsärzte, Betriebsräte etc. verstehbar, annehmbar und machbar erscheinen.

Evaluationen dienen der Transparenz in Projekten und dokumentieren die Wirkungen von Interventionen. Sie können Zusammenhänge aufzeigen. Ziele von Evaluationen sind nach DGUV (2014)

- *Erkenntnisgewinn*: Herausfinden, ob neu definierte oder modifizierte Prozesse oder Interventionen für eine Zielerreichung wirksam sind.
- *Optimierung*: Maßnahmen anpassen oder modifizieren, um die angestrebten Ziele zu erreichen.
- *Wirksamkeitsmessung*: Effektivität der intendierten Maßnahmen belegen.
- *Kontrolle*: Die Durchführung nachgewiesener wirksamer Maßnahmen überwachen
- *Legitimation*: Maßnahmen über den Nachweis ihrer Effektivität legitimieren.

Rossi, Freeman und Hofmann (1988, S.3) bezeichnen Evaluationsforschung als die „systematische Anwendung sozialwissenschaftlicher Forschungsmethoden zur Beurteilung der Konzeption, Ausgestaltung, Umsetzung und des Nutzens sozialer Interventionsprogramme“, welche auf die „Verbesserung der Planung und laufenden Überwachung sowie zur Bestimmung der Effektivität und Effizienz von Gesundheits- [...] maßnahmen“ abzielt. Neben der Überprüfung von Endergebnissen einer Maßnahme (summative Evaluation) kann auch der Verlauf von Intervention durch Evaluationen mitverfolgt und beeinflusst werden (formative Evaluation) (Bortz & Döring, 2006, S.726).

Evaluationen können nicht nur als rückblickende Wirkungskontrolle im Sinne von Vorher-Nachher-Vergleichen durchgeführt werden, sondern auch vorausschauend und prognostisch Veränderungen bewerten. Dabei beruhen die erzielten Ergebnisse und Schlussfolgerungen auf empirisch gewonnenen, beobachteten und/oder gemessenen Daten. Anspruch ist die Nachvollziehbarkeit und Reproduzierbarkeit der Bewertung.

Leider glaubt man oft, der Schwerpunkt von Evaluationen läge nur in der Wirkungskontrolle, die mittels einer Vorher-Nachher-Messung anhand von definierten Kriterien eine Wirkungsmessung

durchführt und danach den Erfolg von Maßnahmen bewertet (Bortz & Döring, 2006; Gollwitzer und Jäger, 2007). Evaluation wird damit sehr verkürzt dargestellt. Es geht zwar um Wirkungskontrolle – aber nicht in der Hauptsache: Hauptzweck von Evaluation ist nicht das Überprüfen und Beweisen der Wirksamkeit von Maßnahmen, sondern ein kontinuierlicher Verbesserungsprozess zum Wohl der Beschäftigten und der Organisation: „Not to prove, but to improve“ wie dies Stufflebeam (2002) auf den Punkt bringt.

Evaluation soll demnach Antworten geben auf die Fragen:

- Tun wir die richtigen Dinge?
- Tun wir die Dinge richtig?

Der Expertenkreis „Evaluation, Wirksamkeit und Nachhaltigkeit“ des Fachverbands Psychologie für Arbeitssicherheit und Gesundheit e.V. (FV PASiG) hat einige Evaluationsmodelle gesichtet und ausgewertet und hält das CIPP-Modell der Programmevaluation von Daniel Stufflebeam (2007) für eine besonders geeignete konzeptionelle Grundlage von Evaluationen.

Das CIPP Modell basiert auf einem umfassenden Verständnis von Evaluation unter Anwendung der Evaluationsstandards des Joint Committee on Standards for Educational Evaluation. Demnach wird Evaluation als die „systematische Untersuchung eines Gegenstandes hinsichtlich seiner Güte, seines Wertes und seiner Bedeutsamkeit“ definiert (Stufflebeam & Shinkfield, 2007, S.326). Darüber hinaus wird dem Modell folgende Definition zugrundegelegt: „Operationally, evaluation is the process of delineating, obtaining, reporting, and applying descriptive and judgmental information about some object’s merit, worth, significance, and probity in order to guide decision making, support accountability, disseminate effective practices, and increase understanding of the involved phenomena” (Stufflebeam & Shinkfield, 2007, S.326).

Das Modell bietet eine überschaubare Systematik, in die sich alle erfolgskritischen Aspekte des Handelns für Arbeitssicherheit und Gesundheit einordnen lassen. Von der Bedarfsermittlung und Zielsetzung bis zur Umsetzung und dem Transfer auf andere Anwendungsfelder können alle Phasen einer Maßnahme begleitet und optimiert werden.

1.2 Das CIPP-Modell

Das CIPP Modell (Stufflebeam, 2007) besteht aus 4 Phasen, die entsprechend dem Akronym die Evaluation des Kontextes (C), des Inputs (I), des Prozesses (P) und des Produkts (P) zum Gegenstand haben. Die einzelnen Phasen dienen folgenden Zielen und Zwecken:

1.2.1 Kontextevaluation (C): „Was muss getan werden?“

Kontextevaluation dient der Zielfindung, der Bedarfsklärung, der Klärung von Ansprüchen, Ressourcen und Problemlagen, der Bestimmung der Voraussetzungen für Maßnahmen, Programme und Projekte. Bei der Kontextevaluation wird überprüft, wer die Nutznießenden einer Maßnahme sind (z.B. Beschäftigte im Betrieb, das Management etc.) und welche Bedürfnisse, Probleme, aber auch Stärken diese Nutznießenden haben (z. B. Wissen, Motivation). Notwendige Aktivitäten hierzu sind Stakeholderanalysen. Die Ergebnisse der Kontextevaluation versetzen den Auftraggeber in die Lage, die Ziele der Maßnahme bzw. des Programms genauer zu verstehen und ggf. zu revidieren.

Das Ergebnis der Kontextevaluation bedeutet: die Ziele und die Rahmenbedingungen für das Erreichen einer Evaluation sind geklärt.

1.2.2 Inputevaluation (I): „Wie muss es getan werden?“

Die Inputevaluation dient dem Beurteilen von Strategien, Konzepten, Ablaufplänen, Vorgehensweisen, der Bestimmung der Planungsqualität und dem Abwägen von Verfahrensalternativen. Bei der Inputevaluation wird systematisch und unter Informationssuche die Entscheidung zwischen verschiedenen Verfahrensplänen erarbeitet. In diesem Zusammenhang wird die Ressourcenverteilung (z.B. auf Trainings und Informationsmaterialien) betrachtet und alternative Verfahrenspläne werden in Hinblick darauf analysiert, ob sie ausreichend, durchführbar und wirtschaftlich sind. Notwendige Aktivitäten hierzu sind die Bewertung von Strategien zur Durchführung und Umsetzung von Interventionen, Maßnahmen, Programmen und Projekten, die Bewertung von Vorgehensweisen, methodischen Wegen und von Ablaufplänen.

Das Ergebnis der Inputevaluation heißt: Es werden die Pläne zur Durchführung und Umsetzung von Interventionen, Maßnahmen, Programmen und Projekten verbessert.

1.2.3 Prozessevaluation (P): „Wurde das Programm/ das Verfahren wie vorgesehen umgesetzt?“

Die Prozessevaluation überwacht, dokumentiert und erfasst Programmaktivitäten und vergleicht sie mit den Zielen des Programms. Durch die kontinuierliche Kontrolle ist es möglich, Pläne und Verfahren zu verbessern und Ursachen für einen Erfolg oder Misserfolg von Maßnahmen oder Interventionen zu identifizieren.

Das Ergebnis der Prozessevaluation ist: Es können die Umsetzungsprozesse und die Durchführung einzelner Maßnahmen oder Interventionen bewertet oder verbessert werden.

1.2.4 Produktevaluation (P): „Führt das Programm zum Erfolg?“

Die Produktevaluation beinhaltet die vier Teilaspekte Einflussevaluation, Effektivitätsevaluation, Nachhaltigkeitsevaluation und Transferevaluation.

- 1) Die Einflussevaluation befasst sich damit, ob die Zielgruppe erreicht wurde. Es wird geprüft, ob die Zielgruppe die Maßnahme erhalten hat, ob sie gerne teilgenommen hat und wie sie die Maßnahme einschätzt.
- 2) Die Effektivitätsevaluation ist das, was die meisten Theoretiker und Praktiker in einer engeren Definition unter einer Evaluation verstehen, nämlich Effektivitätsprüfung der Maßnahmen. Sie bedient sich statistischer Methoden, um z.B. die Wirkung der Maßnahme auf die Zielgruppe und eine Vergleichsgruppe ohne Maßnahme zu vergleichen und eine Aussage treffen zu können.
- 3) Die Nachhaltigkeitsevaluation überprüft schließlich, wie sinnvoll die Einführung einer Maßnahme über einen längeren Zeitraum ist. Es können auch mehrere eingeführte Maßnahmen über einen längeren Zeitraum verglichen werden.
- 4) Die Transferevaluation erfasst Umstände, Verhaltensweisen oder Faktoren, die den Transfer begünstigen, z. B. Erfolgs- und Misserfolgserwartungen in Bezug auf den Transfer, Überzeugungen zu den Erwartungen der Kollegen und Vorgesetzten, eigene Kontrollüberzeugungen. Auch die Übertragbarkeit des Vorgehens oder der Maßnahme auf andere inhaltliche Bereiche, z. B. andere Seminartypen, sind Teil der Transferevaluation.

Das Ergebnis heißt: Man hat die Zielgruppe (mehr oder weniger) erreicht, die Ziele wurden (mehr oder weniger) erreicht, die Maßnahme kann (mehr oder weniger sinnvoll) verstetigt und auf andere Bereiche übertragen werden.

1.2.5 Fazit

Evaluation umfasst also bei weitem nicht nur die (quasi-experimentelle) Wirksamkeitsprüfung nach erfolgter Intervention, sondern auch frühere und qualitative Stufen der rationalen und systematischen Beurteilung von Projekten, Programmen und Maßnahmen. Zum CIPP-Modell liegen zahlreiche Literaturstellen, Anwendungsbeispiele und ausgearbeitete Checklisten für jede Phase von Evaluationsprojekten, sowie für die Metaevaluation vor.

Im weiteren Text folgen Beiträge, welche Anwendungen des CIPP-Modells in der Praxis der Arbeitssicherheit und des Gesundheitsschutzes verdeutlichen sollen. Sie beruhen auf den Beiträgen der Autoren (in Zusammenarbeit mit anderen Kolleginnen) zum 19. Workshop „Psychologie der Arbeitssicherheit und Gesundheit“ im Arbeitskreis „Evaluation: Tun wir die richtigen Dinge? Tun wir die Dinge richtig?“

2 Evaluation? Wo steht das im Gesetz?¹

Anlässe zur gesetzlich geforderten Wirksamkeitsüberprüfung werden in der betrieblichen Arbeitsschutzpraxis häufig übersehen. Wenn Evaluation überhaupt stattfindet, wird sie oft auf die nachträgliche Berechnung wenig aussagekräftiger Kennwerte, etwa Unfallzahlen und Krankenstand, reduziert. In diesem Abschnitt soll aufgezeigt werden, dass das Arbeitsschutzgesetz (ArbSchG, 1996) einen Evaluationsauftrag enthält, der mit Hilfe bewährter Evaluationsmodelle erfüllbar ist.

2.1.1 Arbeitsschutzgesetz: Sicherheit und Gesundheit als Managementprogramm

Ziel und Grundanliegen des betrieblichen Arbeitsschutzes sind „Sicherheit und Gesundheitsschutz der Beschäftigten bei der Arbeit“ (§ 1 Abs. 1 ArbSchG). Dem Arbeitgeber werden im Gesetz (§§ 3 ff.) die Verfahrensregeln vorgegeben, nach denen er eine „umfassende, präventionsorientierte Handlungspflicht“ (Kohte, 2011, S. 150 ff.) zu erfüllen hat.

Um zu „ermitteln, welche [Schutzmaßnahmen] erforderlich sind“ (§ 5 Abs. 1), muss der Arbeitgeber zunächst die Gefährdungen im Betrieb beurteilen. Er hat dann „die erforderlichen Maßnahmen“ (§ 3 Abs. 1) zu planen (§ 3 Abs. 2), „zu treffen, ... auf ihre Wirksamkeit zu überprüfen und erforderlichenfalls sich ändernden Gegebenheiten anzupassen. Dabei hat er eine Verbesserung von Sicherheit und Gesundheitsschutz ... anzustreben“ (§ 3 Abs. 1). Die „festgelegten Maßnahmen ... und das Ergebnis ihrer Überprüfung“ müssen aus „Unterlagen ... ersichtlich“ sein (§ 6 Abs. 1). „Zur

¹ Dieser Abschnitt ist eine geringfügig überarbeitete Version des Beitrags: Watzl, F. (2016). Programmevaluation? Wo steht das im Gesetz? In R. Wieland, K. Sailer & M. Hammes (Hrsg.), Psychologie der Arbeitssicherheit und Gesundheit. Dialog statt Monolog (S. 79-82). Reprint mit der freundlichen Genehmigung des Asanger Verlags.

Planung und Durchführung der Maßnahmen ... hat der Arbeitgeber ... für eine geeignete Organisation zu sorgen und die erforderlichen Mittel bereitzustellen“ (§ 3 Abs. 2). Die Maßnahmen sind „bei allen Tätigkeiten und eingebunden in die betrieblichen Führungsstrukturen“ (§ 3 Abs. 2 Nr. 2) zu beachten. Bereits bei der Arbeitsgestaltung (§ 4 Nr. 1) sind „Technik, Arbeitsorganisation, ... Arbeitsbedingungen, soziale Beziehungen und [Umwelteinflüsse] sachgerecht zu verknüpfen“ (§ 4 Nr. 4).

Das Gesetz erteilt damit ausdrücklich den Auftrag, die Wirksamkeit von Maßnahmen der Arbeitssicherheit und des Gesundheitsschutzes zu überprüfen. Es diktiert aber nicht, *welche* Schutzmaßnahmen im Betrieb zu ergreifen sind und *wie* ihre Wirksamkeit überprüft werden soll. Hinter der fragmentierten Darstellung einzelner Pflichten verbirgt sich vielmehr die Aufforderung, den komplexen, zielgerichteten „Kernprozess einer Vielzahl von Managementsystemen“ zu initiieren, zu steuern und zu optimieren (Elke, Gurt, Möltner & Externbrink, 2015, S. 99).

Gelegenheiten zu handlungsleitender Evaluation bieten sich in allen Phasen dieses Kernprozesses, d.h. prospektiv (ArbSchG: „zu planen“), prozessbegleitend („anzupassen“, „zu verbessern“) und retrospektiv („zu überprüfen“). Modelle der Programmevaluation nutzen diese Ansatzpunkte. Sie ermöglichen damit die Integration einer „Systemkontrolle“ (Kohte, 2011, S. 156), die dem Managementauftrag des Arbeitsschutzgesetzes besser entspricht als die isolierte Erhebung vermeintlicher Erfolgskennzahlen.

2.1.2 Begleitende, wissensbasierte Optimierung: Programmevaluation

Attkisson und Broskowski (1978; zitiert nach Holling, 2009) definieren Programmevaluationen als „vernunftgeleitete Beurteilungen eines Programms hinsichtlich Aufwand, Effektivität und Angemessenheit auf der Grundlage systematischer Datenerhebung und Datenanalyse“ (S. 3). Dabei ist die Aufgabe der Evaluierenden, „durch Bereitstellung relevanten Handlungswissens die Programmsteuerung und -planung zu unterstützen, die Wirkungen eines vorab definierten Programms zu analysieren sowie Erfolgsfaktoren und Hindernisse für die Etablierung und Fortsetzung des Programms zu identifizieren“ (Gollwitzer & Jäger, 2009, S. 14). Phasenmodelle der Programmevaluation koppeln Evaluationsschritte an den Verlauf des Interventionszyklus (Döring, 2009, S. 103).

Das CIPP-Modell der Programmevaluation (Stufflebeam, 2003b; Stufflebeam & Shinkfield, 2007) ist ein solches Phasenmodell, das sich in unterschiedlichsten Anwendungsbereichen bewährt hat. Es betrachtet alle Phasen einer Intervention und eignet sich deshalb auch zur Begleitung und Überprüfung des gesetzlich geforderten Managementprozesses für Sicherheit und Gesundheit bei der Arbeit.

2.1.3 Kontextevaluation: Was soll erreicht werden?

Das globale Ziel „Sicherheit und Gesundheitsschutz“ muss im Betrieb näher spezifiziert werden. Die Bedürfnisse aller Anspruchsgruppen (Arbeitgeber, Beschäftigte, Führungskräfte usw.) werden dabei ebenso berücksichtigt wie die nutzbaren personalen und organisationalen Ressourcen. Auch Hindernisse, die der Zielerreichung im Wege stehen könnten, werden bereits in diesem frühen Stadium bedacht. Schon die Zielformulierung bestimmt, welche Wirkung bestenfalls erreicht werden kann und mit welchen Indikatoren sie messbar sein wird. Die Zielspezifikation richtet den Managementprozess auf die Grundwerte „Gesundheit“ und „Sicherheit“ aus (§ 1 Abs. 1 ArbSchG). Indikatoren für Prozessgüte und Zielerreichung sind eine notwendige Voraussetzung der Wirksamkeitsprüfung nach § 3 ArbSchG. Ziel muss z.B. nicht immer eine Reduktion der Unfälle sein, wenn das Unternehmen hier schon gut Ergebnisse erzielt hat. Dagegen kann man auch Ziele zur Verminderung der psychischen Belastungen und Beanspruchungen, oder für eine verbesserte Zusammenarbeit der Akteure für Sicherheit und Gesundheit bei der Arbeit definieren. Zudem stellt sich die Frage, wie die wesentliche Anspruchsgruppen eingebunden werden sollten (dazu auch Abschnitt 3).

2.1.4 Inputevaluation: Wie soll es erreicht werden?

Ein spezifisches Ziel kann mit verschiedenen Mitteln erreicht werden. Bewertung und Auswahl alternativer Maßnahmen beeinflussen die erreichbaren Wirkungen. In dieser Phase wird die Maßnahmenhierarchie des § 4 ArbSchG zu berücksichtigen sein. Diese meint, dass technische vor organisatorischen und vor personenorientierten Maßnahmen erfolgen sollen.

2.1.5 Prozessevaluation: Wird es getan?

Arbeitsschutzmaßnahmen wirken unter wechselnden inner- und außerbetrieblichen Randbedingungen. Eine wesentliche Einflussgröße ist die jeweilige Sicherheits- und Gesundheitskultur in den Unternehmen. Ob die Maßnahmen unter den konkret gegebenen Umständen planmäßig umgesetzt werden konnten, muss gezielt überprüft werden. So können Bedingungen ihrer Wirksamkeit (oder Wirkungslosigkeit) identifiziert werden. Damit werden verhältnispräventive Ansatzpunkte für die nach § 3 ArbSchG anzustrebende Verbesserung aufgedeckt.

2.1.6 Produktevaluation: Hat es gewirkt?

Stufflebeam (2007) differenziert diese Phase in vier Aspekte. Die Maßnahmen müssen ihre Zielgruppe erreichen (Einflussevaluation). Dies ist jedoch keine hinreichende Voraussetzung ihrer Wirksamkeit. Deshalb müssen die tatsächlichen Effekte registriert, auf das spezifische Ziel bezogen und dokumentiert werden (Effektivitätsevaluation). Ob und wie sich die momentan gemessenen Wirkungen verstetigen lassen, betrachtet die Nachhaltigkeitsevaluation. Dabei darf nicht angenommen werden, dass Maßnahmen, die sich unter bestimmten Bedingungen bewährt haben, ohne Anpassungen

auf andere Zeitpunkte oder Situationen übertragbar sind (Transferevaluation). Ihre Wirkungen sind „zu sichern“, wie auch in § 1 Abs. 1 ArbSchG gefordert.

2.1.7 Fazit

Die Wirksamkeit des Arbeitsschutzmanagements kann in allen Programmphasen begünstigt oder beeinträchtigt werden. Programmevaluation wird im Arbeitsschutzgesetz zwar – unter dieser Bezeichnung – nicht ausdrücklich verlangt. Der umfassende Managementansatz des Gesetzes erfordert jedoch einen ebenso umfassenden Ansatz der Wirksamkeitsprüfung, der über die bloße Registrierung von Effekten hinausgehen muss. Evaluation kann alle Phasen des Arbeitsschutzprozesses unterstützen – von der Bedürfnisermittlung bis zur Verstetigung.

Das CIPP-Modell der Programmevaluation bietet dafür eine vielfach bewährte, praxisnahe Struktur. Checklisten sind ein fester Bestandteil des CIPP-Modells und zugleich in der betrieblichen Praxis gut akzeptiert. Die Entwicklung und Verbreitung arbeitsschutzspezifischer Checklisten zur Programmevaluation ist deshalb zu wünschen. Sie würden etablierte Standards der Evaluationsforschung für betriebliche Anwender verständlich, annehmbar und nutzbar machen. Dies würde zu einem reiferen Verständnis von Wirksamkeit und Wirksamkeitsüberprüfung beitragen.

3 Stakeholder-Analysen als Voraussetzung für gelingende Evaluationen

3.1 Einführung

Der Einsatz von Ressourcen für Arbeitssicherheit und Gesundheitsschutz braucht starke Argumente und Akzeptanz, um sich gegen andere Investitionen in den Unternehmen behaupten zu können. Dabei müssen in allen Bereichen mikropolitische Interessen verschiedener an Arbeitssicherheit und Gesundheit beteiligter Akteure balanciert werden. Maßnahmen für den Arbeits- und Gesundheitsschutz können u.a. mit Arbeitgeber-Interessen, gewerkschaftlichen Interessen und den Interessen der Beschäftigten selbst in Konflikt geraten. Daher muss eine Evaluation gut durch eine Stakeholder-Analyse vorbereitet werden. Stakeholder-Analysen sind damit ein wesentlicher Teil der Kontext-Analyse.

3.2 Stakeholder – Definition

Crane und Matten (2010) nennen Stakeholder (Anspruchsgruppen) diejenigen, die von einer Organisation einen Nutzen oder Schaden zu erwarten haben, oder deren Rechte durch eine Organisation verletzt werden können, oder die Personengruppen, die durch die Organisation respektiert werden müssen. Anspruchsgruppen einer Organisation sind folglich Arbeitgeber selbst, Beschäftigte (und deren Familien), Zulieferer, Kunden, Wettbewerber, Medien, Gesellschaft u.v.m. Organisationen befinden sich in Netzwerken von direkten und indirekten Anspruchsgruppen (Carroll & Buchholtz, 2003, Crane und Matten, 2010). Anspruchsgruppen sind alle internen und externen Personengruppen, die in irgendeiner Form von unternehmerischen Tätigkeiten aktuell oder in Zukunft betroffen sind.

Nach Wheeler und Sillanpää (1997) werden primäre und sekundäre Stakeholder unterschieden: die primären Stakeholder sind direkt von Unternehmenstätigkeiten betroffen, die sekundären Stakeholder haben Einfluss auf die Organisation, sind aber meist nicht direkt von den Auswirkungen betroffen (Carroll & Buchholtz, 2003).

Anspruchsgruppen für Arbeitssicherheit und Gesundheitsschutz sind überwiegend innerhalb des Unternehmen zu finden: Direkt betroffen sind z.B. Beschäftigte, Führungskräfte, Management, Betriebsräte, Betriebsärzte, Fachkräfte für Arbeitssicherheit, Sicherheitsbeauftragte, Arbeitspsychologen, Gesundheitsmanagement, Personalmanagement; betroffen sind möglicherweise aber auch extern die Familienangehörigen. Hinzu kommen z.B. auch Beratende aus den Berufsgenossenschaften, die z.B. die Einhaltung gesetzlicher Vorgaben fordern; oder Kunden, die ggf. Maßnahmen für Sicherheit und Gesundheit bei der Arbeit auf einer Baustelle als überflüssig und zu teuer bewerten.

3.3 Einflussmöglichkeiten verschiedener Anspruchsgruppen in Arbeitssicherheit und Gesundheitsschutz

Nach Mitchell, Agle und Wood (1997) sind die Stakeholder nach ihrem Einfluss, der Legitimität ihrer Ansprüche, und ihrer Dringlichkeit in ihren Ansprüchen zu unterscheiden (Carroll & Buchholtz, 2003). Die verschiedenen Stakeholdermodelle können danach unterschieden werden, ob sie deskriptiv, instrumentell oder normativ sind (Donaldson und Preston, 1995). Sie sind deskriptiv, wenn sie die Art der vorhandenen Stakeholder beschreiben, mit denen eine Organisation(einheit) interagiert; sie sind instrumentell, wenn es darum geht, mit einem Stakeholdermanagement Nutzen für eine Organisation(einheit) zu erzeugen; sie sind normativ, wenn sie die Legitimität der Ansprüche überprüfen (Carroll & Buchholtz, 2003, Crane & Matten, 2010). Im Arbeitsschutz sind mindestens die instrumentelle und die normative Perspektive zu berücksichtigen.

Für ein systematisches Einbinden von Anspruchsgruppen ist es sinnvoll, mehrere Fragen zu beantworten (nach Carroll & Buchholtz, 2003).

- *Wer sind die Anspruchsgruppen?* Für den Bereich Arbeitssicherheit und Gesundheit sind dies Beschäftigte, Führungskräfte, Management, Betriebsräte, Betriebsärzte, Fachkräfte für Arbeitssicherheit, Sicherheitsbeauftragte, Arbeitspsychologen, Gesundheitsmanagement, Personalmanagement, Controlling, Aufsichtspersonen der Berufsgenossenschaft
- *Was sind die Interessen der Anspruchsgruppen?* Hier spielen persönliche Interessen wie Status oder Bequemlichkeit, ökonomische oder politische Interessen eine Rolle, aber auch Rechtsicherheit oder übergeordnete Werteorientierungen, z.B. Sicherheit unter allen Umständen zu gewährleisten (Vision Zero) oder den Auftrag des Kunden unter allen Umständen fertig zu stellen.
- *Welches Machtpotenzial haben die Anspruchsgruppen?* Machtausübung mit negativer Ausprägung ist möglich durch eine Verhinderung von Maßnahmen durch Vetorechte oder fehlende Finanzierung, fehlendes Commitment und Arbeit nach Vorschrift; eine unterstützende Form meint Zugang zu wichtigen Informationen und das Bereitstellen finanzieller und zeitlicher Ressourcen.
- *Welche Verantwortlichkeiten bestehen gegenüber Anspruchsgruppen?* Bewertet wird, wie legitim deren Ansprüche sind, z.B. Bestehen rechtliche Vorgaben aus dem Arbeitsschutzgesetz? Fallen die Ansprüche in den Bereich der moralisch-sittlichen Normen (z.B. Themen der Work-Life-Balance)? Oder sind manche Ansprüche ggf. nur vollständig freiwillige Leistungen einer Organisationseinheit, z.B. auch Privates am Arbeitsplatz erledigen zu können.
- *Welche Maßnahmen sollten ergriffen werden?* Einbindung der Anspruchsgruppen durch Formen der Partizipation, Information, Entscheidung, ...

Man mag sich nun als ein konkretes Beispiel vorstellen, dass Maßnahmen zur Verringerung psychischer Belastungen und Beanspruchungen evaluiert werden sollen. So können z.B. Interessen darin liegen, dass eigentlich unwirksame Trainings-Maßnahmen im Ergebnis positiv bewertet werden sollen, da (a) sie kostengünstige Pseudo-Maßnahmen sind und man dann die Prozesse nicht weiter reflektieren muss; (b) bestimmte Trainer sich im Unternehmen etabliert haben, (c) Führungskräfte ihren Führungsstil nicht überdenken wollen (vergleiche Tabelle 1 nächste Seite).

Ein anderes Beispiel könnte die Einführung von computergestützten Unterweisungen sein, die höchst unterschiedliche Reaktionen bei den vielen Betroffenen hervorrufen können.

Diese Reaktionen, oder besser die dahinter liegenden Interessen und Bedürfnisse haben Auswirkungen auf die Zielsetzung von Evaluationen und darauf, welche Indikatoren zu welchen Zeitpunkten erhoben werden. Die Praxis zeigt: Systematische Vorher-Nachher-Messungen um eine Intervention zu bewerten, unterbleiben oft. Kritische Variablen werden gar nicht erhoben.

Im nächsten Schritt interessiert, wie sehr die einzelnen Anspruchsgruppen ihre Interessen durchsetzen können, und ob eine Organisation diese Interessen als berechtigt berücksichtigen sollte um daraus Maßnahmen der Partizipation abzuleiten. Unterschiedliche Partizipationsgrade sollten hier Berücksichtigung finden (Tabelle 2, Seite 14).

3.4 Fazit

Erst nach einer ausführlichen Stakeholder-Analyse wird deutlich, welche unterschiedlichen Werte, Interessen und Motive im Prozess berücksichtigt werden müssen. Erst durch eine derartige Analyse lassen sich Umsetzungsschwierigkeiten vorhersehen und entsprechende Maßnahmen planen, was die Erfolgsaussichten für eine Evaluation deutlich erhöht. Deutlich wird auch, dass im Vorfeld umfangreiche Maßnahmen zur Kommunikation und Partizipation erfolgen müssen. Die interne Kommunikation und Abstimmung mit verschiedenen Betroffenen erst kann ein Gelingen vorbereiten.

Ist der Kontext der Evaluation geklärt, erfolgt bei der Input-Evaluation die Entscheidung zwischen alternativen Verfahrenswegen in Hinblick auf Hinlänglichkeit, Durchführbarkeit und Wirtschaftlichkeit. Danach geht es an die Umsetzung des Projektes und die begleitende Prozessevaluation.

Tabelle 1: Mögliche Anspruchsgruppen und ihre Interessen bei einer möglichen Evaluation von Maßnahmen zur Verringerung psychischer Belastungen und Beanspruchungen

Anspruchsgruppe	Interessen	Mögliche Folgen
Führungskräfte	<ul style="list-style-type: none"> – Persönlich: positives Selbstbild erhalten, Karriereoptionen erhalten – Teamorientiert: Hohe Arbeitsleistung und Zufriedenheit bei den Beschäftigten erhalten – Für die Organisation: zusätzliche Anstrengungen vermeiden / Ressourceneinsatz schonen 	<ul style="list-style-type: none"> – Wohlfühl-Pakete werden ggf. gerne akzeptiert – Scheinergebnisse können erzeugt und nach oben kommuniziert werden – Finanzielle Aspekte werden in den Vordergrund gerückt
Beschäftigte	<ul style="list-style-type: none"> – Persönlich: Arbeitslast reduzieren, Sinnvolle Arbeit erleben, zusätzliche Anstrengungen vermeiden – Für die Organisation: gute Arbeit leisten 	<ul style="list-style-type: none"> – Hilfen jeder Art werden gerne angenommen – Trainings sind ggf. willkommene „Auszeit“ – Umstrukturierungen haben einen negativen Beigeschmack
Controlling	<ul style="list-style-type: none"> – Für die Organisation: Kosten reduzieren 	<ul style="list-style-type: none"> – nur am finanziellen Bereich interessiert: Unzureichende Indikatoren und KPIs für Unternehmenssteuerung liegen vor
Externe Trainer	<ul style="list-style-type: none"> – Persönlich/ ihre Organisation: Aufträge generieren und sicherstellen 	<ul style="list-style-type: none"> – „Wohlfühl-Pakete“ werden angeboten, konfliktträchtige Bereiche werden eher vermieden
Betriebsrat	<ul style="list-style-type: none"> – Persönlich: Macht ausüben – Teamorientiert: Hohe Zufriedenheit bei den Beschäftigten erhalten – Gewerkschaftlich: Position der Beschäftigten stärken 	<ul style="list-style-type: none"> – Befürwortet Maßnahmen, verteidigt aber eher ökonomische Interessen der Beschäftigten
....	– ...	– ...

Tabelle 2: Stakeholder-Analyse mit Maßnahmen für eine Evaluation von Maßnahmen

Anspruchsgruppe	Interessen	Macht-potenzial	Legitimität des Anspruchs	Maßnahmen zur Einbindung
Führungskräfte	<ul style="list-style-type: none"> – Persönlich: positives Selbstbild erhalten, Karriereoptionen erhalten – Teamorientiert: Hohe Arbeitsleistung und Zufriedenheit bei den Beschäftigten erhalten – Für die Organisation: zusätzliche Anstrengungen vermeiden / Ressourceneinsatz schonen 	hoch	<p>mittel</p> <p>hoch</p> <p>hoch</p>	Frühzeitig einbinden, um Akzeptanz sicherzustellen, Umfangreiche Informationen bereitstellen
Beschäftigte	<ul style="list-style-type: none"> – Persönlich: Arbeitslast reduzieren, Sinnvolle Arbeit erleben, zusätzliche Anstrengungen vermeiden – Für die Organisation: gute Arbeit leisten 	gering bis mittel (je nach Organisationsgrad)	hoch	Frühzeitig Informieren, Anhörungsrechte,
Controlling	<ul style="list-style-type: none"> – Für die Organisation: Kosten reduzieren 	hoch	hoch (Fortbestand der Organisation sichern)	Umfangreiche Informationen bereitstellen, frühzeitig Informieren
Externe Trainer	<ul style="list-style-type: none"> – Persönlich/ ihre Organisation: Aufträge sicherstellen 	Gering (abhängige Position)	Mittel (Planungssicherheit)	frühzeitig Informieren
Betriebsrat	<ul style="list-style-type: none"> – Persönlich: Macht ausüben – Teamorientiert: Hohe Zufriedenheit bei den Beschäftigten erhalten – Gewerkschaftlich: Position der Beschäftigten stärken 	hoch (Mitbestimmungs- und Vetorechte)	mittel	Umfangreiche Informationen bereitstellen, frühzeitig Informieren, aktiv einbinden

4 Prozessevaluation eines Projektes zum betrieblichen Gesundheitsmanagement²

4.1 Gesundheits-Balanced-Scorecard

In einem Pilotprojekt der Technischen Universität Darmstadt und einer Klinik-Gesellschaft wurde eine Gesundheits-Balanced-Scorecard (Gesundheits-BSC) zur strategischen Steuerung des betrieblichen Gesundheitsmanagements (BGM) implementiert und evaluiert. Die Balanced-Scorecard wird im Kontext des BGM als Steuerungs- und Messkonzept zur unternehmensinternen Koordination von Gesundheitsmaßnahmen verstanden. Dabei werden in einer Gesundheits-BSC die strategischen Ziele des BGM operationalisiert, indem konkrete Maßnahmen abgeleitet werden und deren Zielerreichung überprüft wird (Horváth et al., 2009). Das Pilotprojekt wurde in den Jahren 2012 bis 2014 in zwei Kliniken, die hinreichend vergleichbar waren, durchgeführt. In einem quasi-experimentellen Evaluationsplan diente dabei die eine Klinik als Interventionsklinik und die andere als Warte-Kontrollklinik.

Die Evaluation solcher Projekte erfolgt in den meisten Fällen durch die Methoden der Produktevaluation und trifft Aussagen über deren Wirkung bzw. Erfolg nach Abschluss des Projekts (Saunders, Evans & Joshi, 2005). Rein summativ ausgerichtete Herangehensweisen reichen jedoch nicht aus, um komplexe Programme unter realen Bedingungen zu evaluieren (Stufflebeam & Shinkfield, 2007). Deshalb wurde das CIPP-Modell von Daniel Stufflebeam als Grundlage für die Evaluation des gesamten Pilot-Projekts gewählt. Dieses Modell ist ein umfassendes Rahmenmodell zur Anleitung summativer und formativer Evaluationen von Programmen und folgt dem Grundsatz, dass der wichtigste Zweck von Evaluation nicht zu beweisen sondern zu verbessern ist - „the most important purpose of evaluation is not to prove, but to improve“ (Stufflebeam & Shinkfield, 2007, S.331). Analog des CIPP-Modells wurde für die Implementierung der Gesundheits-BSC eine Kontext-, Input-, Prozess-, sowie Produktevaluation durchgeführt. Im Folgenden wird die Prozessevaluation beschrieben.

² Dieser Abschnitt ist eine geringfügig überarbeitete Version des Beitrags: Schwärzel, L. (2016). Prozessevaluation eines Projektes zum betrieblichen Gesundheitsmanagement. In R. Wieland, K. Sailer & M. Hammes (Hrsg.), *Psychologie der Arbeitssicherheit und Gesundheit*. Dialog statt Monolog (S.75-78). Reprint mit der freundlichen Genehmigung des Asanger Verlags.

4.2 Prozessevaluation

Im Rahmen des CIPP-Modells wird unter Prozessevaluation eine fortlaufende Überprüfung der Umsetzung und Dokumentation von Projektaktivitäten sowie deren Abgleich mit den Zielen und dem Zeitplan verstanden. Die Prozessevaluation erfolgt während der Projektlaufzeit und geht der Frage nach „Is it being done?“ (Stufflebeam & Shinkfield, 2007, S.327, ff.). Das formative Vorgehen erfordert eine gewisse Vertrautheit mit dem Projekt sowie einen guten Kontakt zu den Projektbeteiligten, weshalb die Evaluation bestenfalls intern durchgeführt wird (Westermann, 2002). Der Erfolg einer Gesundheits-BSC hängt maßgeblich mit der Qualität ihrer Implementierung zusammen. Das Ziel der Prozessevaluation war es Implementationsfehler zu identifizieren und Strategien zur Prozessoptimierung zu entwickeln. Mit den Auftraggebern wurden 22 explorativ ausgerichtete Evaluationsfragen generiert. Die Beantwortung dieser Evaluationsfragen erfolgte durch unterschiedliche Methoden sowie verschiedene Datenquellen und folgte somit dem Prinzip der Triangulation. Beispielhaft werden im Folgenden zwei Kernevaluationsfragen behandelt:

a) Wurden die Meilensteine des Gesundheits-BSC-Projekts wie geplant erreicht?

Zur Beantwortung dieser Evaluationsfrage erfolgte eine Dokumentenanalyse. Diese wurde durch die Sichtung des Projektplans, der Zwischenberichte sowie aller im Projekt erfassten Sitzungsprotokolle umgesetzt. Es wurden die Projektaktivitäten in Form einer Excel-Tabelle zusammengetragen und ausgewertet. Hierbei wurden die Meilensteine ebenfalls mit den Indikatoren eines BSC-Implementierungsmodells von Horváth und Partners (2007) in Bezug gesetzt, um die Vollständigkeit der Programmimplementierung zu überprüfen.

b) Welche Probleme und Erfolgsfaktoren gab es bei den einzelnen Projektschritten?

Zur Beantwortung der Evaluationsfrage, welche Probleme und Erfolgsfaktoren es bei den einzelnen Projektschritten gab und welche organisationalen Faktoren die Implementierung der Gesundheits-BSC behindert bzw. erleichtert haben, wurden sowohl Experteninterviews durchgeführt als auch ein Fragebogen eingesetzt. Der halbstandardisierte Interviewleitfaden umfasste 26 offene Fragen. Im Fragebogen kamen 11 als Statement formulierte Items zum Einsatz.

4.3 Methodik

Von insgesamt 23 Projektbeteiligten nahmen 15 an der Befragung teil. Davon waren drei Personen für die TU-Darmstadt und 12 für die Interventionsklinik tätig. Von den 15 Personen waren 11 von Beginn des Projekts an beteiligt und vier Personen begannen erst im Verlauf des Projekts mit ihrer Mitarbeit.

Die Auswertung der Experteninterviews erfolgte durch die qualitative Inhaltsanalyse. Dabei wurde die Häufigkeitsanalyse nach Mayring (2010) angewendet. Nach der Bildung von Kategorien, erfolgte die Zuordnung der Textbestandteile zu den jeweiligen Kategorien durch zwei unabhängige Rater. Um die Güte der Intercoderreliabilität zu bestimmen, wurde der Übereinstimmungskoeffizient von Holsti (1969, zitiert nach Mayring, 2010, S.120) herangezogen. Die Fragebogendaten wurden deskriptiv ausgewertet.

4.4 Ergebnisse

Im Rahmen der durchgeführten Prozessevaluation konnte festgestellt werden, dass alle Meilensteine des Gesundheits-BSC-Projekts erreicht wurden. Einige Projektschritte wurden jedoch mit zeitlicher Verzögerung oder gar nicht abgeschlossen. Gründe hierfür waren unter anderem die fehlende Bereitstellung wichtiger Ressourcen, unklare Rahmenbedingungen sowie die mangelnde Unterstützung des Managements. Als Promotoren des Projekts stellten sich das Engagement der Projektteilnehmer/innen, der Einsatz eines Gesundheitsmanagers, die Installation eines Arbeitskreises Gesundheit sowie die wissenschaftliche Unterstützung durch die TU-Darmstadt heraus.

Anhand dieser Erkenntnisse wurden Strategien zur Prozessoptimierung für die Interventionsklinik abgeleitet. Nach einer Priorisierung der Gesundheitsziele, gilt es aufeinander abgestimmte Maßnahmen der Verhaltens- und Verhältnisprävention zu standardisieren und deren Weiterfinanzierung sicherzustellen. Die Einführung von Zielvereinbarungsgesprächen sowie die Weiterbeschäftigung des Gesundheitsmanagers stellen in diesem Prozess weitere wichtige Meilensteine dar. Entscheidend für die vollständige Implementierung der Gesundheits-BSC ist jedoch die Weiterführung des Arbeitskreises Gesundheit, dessen Ausstattung mit ausreichenden Ressourcen und nicht zuletzt die volle Unterstützung des Managements.

4.5 Handlungsempfehlungen

Die Prozessevaluation erlaubt einen Blick in die Black Box eines Programms und liefert Hinweise warum ein Programm gescheitert ist oder erfolgreich war (Linnan & Steckler, 2002). Die Umsetzung der abgeleiteten Strategien zur Prozessoptimierung ist, neben dem systematischen Vorgehen von Diagnose, Intervention und Evaluation, entscheidend für den Erfolg eines strategischen Gesundheitsmanagements (Gurt et al., 2014). Aus den Ergebnissen der Prozessevaluation können praxisrelevante Handlungsempfehlungen für Projekte zur Implementierung einer Gesundheits-BSC ausgesprochen werden. Unternehmen, die ein solches Projekt planen, sollten vorab folgende Punkte berücksichtigen:

- Finanzierungsmöglichkeiten für das BGM eruieren und klare finanzielle Rahmenbedingungen schaffen
- Steuerung des BGM durch die Installation eines Arbeitskreises Gesundheit sowie eines Gesundheitsmanagers sicherstellen
- Unterstützung des Managements sicherstellen
- Strukturen und Prozesse des bisherigen BGM erfassen
- Erhebungsmöglichkeiten von harten Kennzahlen klären und sicherstellen
- Unterstützung durch externe Gesundheits-Experten/ BSC-Experten einholen

Schlussendlich will man bei einer Evaluation natürlich auch wissen, ob man seine Ziele erreicht hat. Dazu führt man eine Produktevaluation durch. Der folgende, letzte Abschnitt beschreibt ein Beispiel einer umfangreichen Produktevaluation eines Projektes zum Gesundheitsmanagement.

5 Produktevaluation am Beispiel „Gesundheits-Balanced-Scorecard“³

5.1 Gesundheits-Balanced-Scorecard

Aufgrund der demographischen Entwicklung, des Fachkräftemangels, steigender Patientenzahlen und der zunehmenden psychischen Belastung von Beschäftigten im Gesundheitswesen besteht akuter Handlungsbedarf für Kliniken, die Gesundheit ihrer Beschäftigte zu erhalten und zu fördern. Im Pilot-Projekt „Gesundheits-Balanced-Scorecard“ (Keil & Vogt, 2012) wurde mit einer Klinikgesellschaft, die vor allem psychiatrische Kliniken betreibt, im Zeitraum von 2012 bis 2015 eine Gesundheits-Balanced-Scorecard (Gesundheits-BSC) zur strategischen Steuerung des betrieblichen Gesundheitsmanagements entwickelt, implementiert und evaluiert. Die Balanced-Scorecard (BSC) ist ein kennzahlengesteuertes Managementsystem (Kaplan & Norton, 1997). Sie übersetzt die Gesundheitsvision in Ziele sowie Kennzahlen und ist dabei in vier Perspektiven unterteilt (Potential-, Prozess-, Beschäftigten- und Erfolgsperspektive). Das Pilotprojekt wurde in zwei Kliniken der Klinikgesellschaft, die hinreichend vergleichbar waren, durchgeführt. In einem quasi-experimentellen Versuchsplan diente dabei die eine Klinik als Interventionsklinik und die andere als Kontrollklinik. Zur Evaluation des Pilotprojekts wurde das CIPP-Modell von Daniel Stufflebeam (2007) herangezogen. Im Folgenden werden insbesondere die Methoden der Produktevaluation beschrieben (siehe auch Keil, 2009).

5.2 Produktevaluation

Die Produktevaluation beschäftigt sich in erster Linie mit der Wirksamkeitsüberprüfung der eingeführten Maßnahmen. Der erste Schritt der Produktevaluation ist die Einflussevaluation. Die zentrale Frage der Einflussevaluation ist, inwiefern die Zielgruppenerreichung stattgefunden hat. Denn die Ergebnisse der darauffolgenden Effektivitätsevaluation können nur dann sinnvoll interpretiert werden, wenn bekannt ist, in welchem Ausmaß die beabsichtigte Zielgruppe erreicht und die Maßnahmen erfolgreich implementiert wurden.

Im ersten Schritt der Einflussevaluation wurde überprüft, inwiefern die im Gesundheits-BSC-Projekt abgeleiteten Maßnahmen angeboten sowie umgesetzt werden und ob im Projekt entwickelte Medien verwendet werden. Im nächsten Schritt wurde überprüft, ob die

³ Dieser Abschnitt ist eine geringfügig überarbeitete Version des Beitrags: Keil, U., Schwärzel, L., Elger, A. & Schneider, K. (2016). Produktevaluation am Beispiel „Gesundheits-Balanced-Scorecard“. In R. Wieland, K. Sailer & M. Hammes (Hrsg.), *Psychologie der Arbeitssicherheit und Gesundheit. Dialog statt Monolog* (S.563-566). Reprint mit der freundlichen Genehmigung des Asanger-Verlags.

Beschäftigten der drei am Projekt beteiligten Stationen die angebotenen Maßnahmen und Materialien überhaupt kennen.

Die summative Effektivitätsevaluation ist der eigentliche Kern der Produktevaluation. Die Wirksamkeit der Gesundheits-BSC wird dabei anhand statistischer Analysen überprüft. Im Rahmen des Pilot-Projekts wurde ein Fragebogen, der die weichen Faktoren der Gesundheits-BSC misst, entwickelt. Darüber hinaus wurden Controllingdaten erhoben, wie z.B. Fehlzeiten und Ausfallkosten. Der Prä-Post-Vergleich zeigt, ob die Zielvorgaben in der Gesundheits-BSC erreicht wurden oder nicht.

5.3 Ergebnisse

Im Untersuchungszeitraum von 2012 bis 2015 konnten 28 von 56 gesetzten Zielwerten der Gesundheits-BSC erreicht werden. 20 Ziele wurden nicht erreicht und bei fünf Kennwerten zeigte sich eine Verschlechterung.

Zu Beginn des Projekts wies die Interventionsklinik verglichen mit der Kontrollklinik geringere Fehlzeitenkosten auf. Betrachtet man die Fehlzeitenquoten beider Kliniken in diesem Zeitraum (07/2011 – 06/2012), lässt sich kaum ein Unterschied feststellen (6,02% vs. 5,89%).

Während der Projektlaufzeit stiegen in beiden Kliniken die fehlzeitenbedingten Kosten an. Auffällig ist die Entwicklung in der Kontrollklinik. Im Zeitraum 07/2012 -06/2013 stiegen die Fehlzeitenkosten fast auf das Doppelte des Ausgangswertes an und sanken im darauffolgenden Zeitraum wieder drastisch ab. Die Fehlzeitenkosten in der Interventionsklinik stiegen im Zeitraum 07/2011 – 06/2012 sowie 07/2012 - 06/2013 ebenfalls an und blieben 07/2013 – 06/2014 nahezu auf dem gleichen Niveau. Beachtlich ist auch die hohe Fluktuationsquote von 17,28% in der Interventionsklinik im Zeitraum 07/2012 – 06/2013. Zu den Langzeiterkrankungen und Arbeitsunfällen liegen aus dem Zeitraum 07/2011 – 06/2012 keine Daten der Kontrollklinik vor. Die Quote der Langzeiterkrankungen ist in der Interventionsklinik insgesamt auf einem höheren Niveau als in der Kontrollklinik.

Anhand der Mitarbeiterbefragungen lassen sich Aussagen zur Wirksamkeit des Gesundheits-BSC-Projekts ableiten. Die Rücklaufquoten zeigen jedoch einen deutlichen Rückgang von über 50% bei der Prä-Erhebung auf 36,9% bei der Post-Erhebung in der Kontrollklinik und auf 23,3% in der Interventionsklinik. Trotzdem lassen sich einige Effekte der Interventionen aufzeigen. Um zu überprüfen, ob sich ein Effekt der Gesundheits-BSC in den Befragungsergebnissen widerspiegelt wurden Varianzanalysen mit Messwiederholung gerechnet. Drei mögliche Ergebnisse können dabei unterschieden werden:

- Haupteffekt der Zeit: Er besagt, dass sich die Mittelwerte beider Kliniken verbessert oder verschlechtert haben, unabhängig von der Einführung der Gesundheits-BSC.
- Haupteffekt der Klinik: Er besagt, dass sich die beiden Kliniken unterscheiden.

- Interaktionseffekt: Er besagt, dass die Veränderungen über die Zeit von der Klinik abhängen. Hat sich der jeweilige Wert in der Interventionsklinik verbessert und ist in der Kontrollklinik in etwa gleich geblieben, kann das so interpretiert werden, dass diese Veränderungen auf die Gesundheits-BSC zurückgeführt werden können.

Von 52 inferenzstatistischen Analysen konnte bei sieben eine signifikante Interaktion bzw. eine Tendenz zur Interaktion festgestellt werden. Im Folgenden werden nur statistisch signifikante Ergebnisse mit einer Irrtumswahrscheinlichkeit kleiner 5% ($p < 0,05$) und solche, die dahin tendieren ($p < 0,1$), berichtet. Die eingeschränkte Leistungsfähigkeit der Beschäftigten der Interventionsklinik nimmt über die Zeit ab. In der Kontrollklinik bleibt sie etwa gleich. Die Familienfreundlichkeit der Interventionsklinik startet auf einem etwas geringeren Niveau als in der Kontrollklinik, holt dann aber auf, wobei sie in der Kontrollklinik etwas sinkt. Die Pause wird von den Beschäftigten der Interventionsklinik über die Zeit hinweg als weniger gefährdet bewertet als in der Kontrollklinik. Die langfristige psychische Beanspruchung nimmt in der Interventionsklinik über die Zeit ab. In der Kontrollklinik bleibt sie in etwa gleich. Der Personalmangel wird von den Beschäftigten der Interventionsklinik über die Zeit als weniger belastend bewertet, in der Kontrollklinik hingegen steigt der Wert etwas. Die physische Beanspruchung nimmt in der Interventionsklinik über die Zeit marginal ab, in der Kontrollklinik steigt sie wiederum von einem geringeren Ausgangsniveau an. Der Präsentismus verringert sich stärker in der Interventionsklinik über die Zeit als in der Kontrollklinik. Die Zufriedenheit mit der Arbeitssituation nimmt in der Interventionsklinik zu, in der Kontrollklinik bleibt sie gleich.

Neben der Einfluss- und Effektivitätsevaluation sind die Nachhaltigkeits- und Transferevaluation weitere Bestandteile der Produktevaluation. Die Nachhaltigkeitsevaluation war nicht mehr Projektbestandteil. Somit können hierzu keine Aussagen getroffen werden. Erst nach einiger Zeit der Steuerung des Gesundheitsmanagements mit Hilfe der Gesundheit-BSC lassen sich belastbare Ergebnisse zur Nachhaltigkeit aufzeigen. Zur Transferevaluation, die sich mit dem Roll-Out der Gesundheits-BSC auf weitere Kliniken beschäftigt, konnten bereits Handlungsempfehlungen abgeleitet werden.

5.4 Diskussion

Diskussion der Methode: Zunächst muss einschränkend gesagt werden, dass die Rücklaufquoten in der Nachbefragung deutlich nachgelassen haben und der echte Vorher-Nachher-Vergleich nur auf einer recht kleinen Stichprobe basiert. Zudem konnten wegen zeitlicher Verzögerungen in den ersten Phasen des Projektes nur wenige Maßnahmen im Projektzeitraum umgesetzt werden, welche zusätzlich zu wenig Zeit hatten, um auf die Belegschaft einzuwirken. Eine vorgeschlagene zeitliche Streckung des Projektes wurde vom Auftraggeber abgelehnt.

Die Einflussevaluation ergab, dass nur zwei der im Gesundheits-BSC-Projekt entwickelten Maßnahmen kontinuierlich angeboten werden. Entsprechend der Konzeption der Maßnahmen kann die Zielgruppenerreichung positiv bewertet werden. Die Einflussnahme des Programms konnte von den Beschäftigten der drei Stationen jedoch nicht beurteilt werden.

Bei der Vorher-Nachher-Befragung in der Interventions- und der Kontrollklinik zeigten sich Zielerreichungen in 28 von 56 Kennwerten und in sieben Fällen Veränderungen im Sinne des quasi-experimentellen Versuchsplans. Allerdings wurden auch in 45 Fällen Ergebnisse gefunden, die keine Wirkung der Intervention anzeigen. Die mehrdeutigen Ergebnisse bezüglich der Controllingdaten zeigen, dass sich wohl andere Einflüsse als die Interventionen auf diese Größen ausgewirkt haben. Hierbei sind vor allem Umstrukturierungen in den Kliniken zu nennen. Trotz allem erbrachte die Evaluation des Projektes zahlreiche wertvolle Informationen, die für die Klinikgesellschaft wichtig waren, um ihr Gesundheitsmanagement nachhaltig zu verbessern.

5.5 Zusammenfassung

Die vorangegangenen Ausführungen und Projektbeispiele sollten einen Einblick in das Themengebiet der Evaluation geben, die Notwendigkeit der Evaluation zeigen und die einzelnen Schritte des Vorgehens anhand von Projektbeispielen verdeutlichen. Sie haben hoffentlich gezeigt, warum Evaluation (z.B. anhand des CIPP-Modells) sinnvoll und praktisch durchführbar ist. Dabei kann man auch einen kleineren Rahmen wählen, als es in den beschriebenen Projekten der Fall war. Die Grundlogik der Evaluation bleibt gleich. Nur wenn wir Maßnahmen der Arbeitssicherheit und des Gesundheitsschutzes evidenzbasiert herleiten und evaluieren, werden wir auf Dauer messbare Fortschritte in diesem Bereich machen.

6 Literatur

- ArbSchG. (1996). Gesetz über die Durchführung der Maßnahmen des Arbeitsschutzes zur Verbesserung der Sicherheit und des Gesundheitsschutzes der Beschäftigten bei der Arbeit (Arbeitsschutzgesetz).
- Bortz, J. & Döring, N. (2009). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Berlin, Heidelberg: Springer.
- Bortz, J. & Döring, N. (2006). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler (4. Aufl.). Heidelberg: Springer.
- Carroll, A. & Buchholtz, A. (2003). Business and Society (5 ed). South Western: Thomson.
- DeGEVal, (2008 a): Standards für Evaluation. On-line in Internet: <http://www.degeval.de/publikationen/standards-fuer-evaluation/>. Abrufdatum 22.11.2016
- DeGEVal, (2008 b): Leitfaden Evaluation für die Wirtschaft. On-line in Internet: http://www.degeval.de/fileadmin/users/Arbeitskreise/AK_Wirtschaft/Leitfaden_Evaluation_Wirtschaft_DeGEval.pdf. Abrufdatum 22.11.2016
- DGUV, Deutsche Gesetzliche Unfallversicherung (2014) Leitpapier zur Evaluation. Grundverständnis in der gesetzlichen Unfallversicherung. DGUV Grundsatz 311-001.
- Donaldson, T. & Preston, L (1995). The stakeholder theory of the corporation: Concepts, Evidence, Implications. *Academy of Management Review*, 20 (1), 65-91.
- Döring, N. (2009). Phasen der Evaluationsforschung. In H. Holling (Hrsg.), *Grundlagen und statistische Methoden der Evaluationsforschung (Enzyklopädie der Psychologie, B-IV-1, S. 99-134)*. Göttingen: Hogrefe.
- Elke, G., Gurt, J., Möltner, H. & Externbrink, K. (2015). Arbeitsschutz und betriebliche Gesundheitsförderung. Vergleichende Analyse der Prädiktoren und Moderatoren guter Praxis. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin. Verfügbar unter www.baua.de/dok/5992164
- Freeman, E. (1994). *Strategic management. A stakeholder Approach*. Boston: Pitman.
- Gabler Wirtschaftslexikon, Stichwort: Anspruchsgruppenverfügbar unter: <http://wirtschaftslexikon.gabler.de/Archiv/1202/anspruchsgruppen-v6.html>
- Gollwitzer, M. & Jäger, R. (2007). *Evaluation. Workbook*. Weinheim: Beltz Psychologie Verlagsunion.
- Gollwitzer, M. & Jäger, R. S. (2009). *Evaluation kompakt*. Weinheim: Beltz.
- Gurt, J., Möltner, H. & Elke, G. (2014). Erfolgsfaktoren im Arbeitsschutz und in der betrieblichen Gesundheitsförderung. In M. Eigenstetter, T. Kunz, R. Portuné & R. Trimpop. (Hrsg.). *Psychologie der Arbeitssicherheit und Gesundheit 18.Workshop 2014.(S.219-222)*. Kröning: Asanger.
- Holling, H. (2009). Grundlagen der Evaluationsforschung. In H. Holling (Hrsg.), *Grundlagen und statistische Methoden der Evaluationsforschung (Enzyklopädie der Psychologie, B-IV-1, S. 1-34)*. Göttingen: Hogrefe.

- Horváth & Partners (Hrsg.). (2007). *Balanced Scorecard umsetzen* (4. Aufl.). Stuttgart: Schaeffer-Poeschel.
- Horváth, P., Gamm, N., Möller, K., Kastner, M., Schmidt, B., Iserloh, B., Kliesch, G., Otte, R., Braun, M., Matter, M., Pennig, S., Vogt, J. & Köper, B. (2009). *Betriebliches Gesundheitsmanagement mit Hilfe der Balanced Scorecard* (Projekt nr.: F 2126). Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.
- Kaplan, R.S. & Norton, D.P. (1997). *Balanced Scorecard. Strategien erfolgreich umsetzen*. Stuttgart: Schäffer-Poeschel.
- Keil, U. & Vogt, J. (2012). *Balanced Scorecard Gesundheit*. In G. Athanassiou, S. Schreiber-Costa & O. Sträter (Hrsg.). *Psychologie der Arbeitssicherheit und Gesundheit 17. Workshop 2012*.(S.383-386) Krönig: Asanger Verlag.
- Keil, U. (2009). *Evaluation im Arbeits- und Gesundheitsschutz im Wandel der Zeit*. In B. Ludborzs, H. Nold (Hrsg.). *Psychologie der Arbeitssicherheit und Gesundheit. Entwicklungen und Visionen 1980 – 2008 – 2020*, Krönig: Asanger S. 92-104.
- Keil, U., Schwärzel, L., Elger, A. & Schneider, K. (2016). *Produktevaluation am Beispiel „Gesundheits-Balanced-Scorecard“*. In Wieland, R., Seiler, K. & Hammes, M. (Hrsg.), *Psychologie der Arbeitssicherheit und Gesundheit. Dialog statt Monolog*. (S.563-566). Krönig: Asanger Verlag.
- Kohte, W. (2011) § 3: *Grundpflichten des Arbeitgebers*. In N. Kollmer & T. Klindt (Hrsg.), *Arbeitsschutzgesetz. Kommentar, 2. Aufl.* München: Beck (S. 146-187).
- Linnan, L. & Steckler, A.(2002). *Process Evaluation for Public Health Interventions and Research: An Overview*. In L. Linnan & A. Steckler (Hrsg.), *Process Evaluation for Public Health Interventions and Research* (S.1-23). San Francisco: Wiley.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse* (11.überarb. Auflage). Weinheim: Beltz.
- Mitchell, R. K., Agle, B. R. & Wood D. J. (1997). *Towards a theory of stakeholder identification and salience. Defining the principles of who and what really counts*. In: *Academy of Management Review*. 4/22, S. 853–886.
- Rossi, P.H., Freeman, H.E. & Hofmann, G. (1988). *Programm-Evaluation Einführung in die Methoden angewandter Sozialforschung*. Stuttgart: Enke.
- Saunders, R.P., Evans, M.H. & Joshi, P. (2005). *Developing a Process-Evaluation Plan for Assessing Health Promotion Program Implementation: A How-To Guide*. *Health Promotion Practice*, 6 (2), 134-147. Doi: 10.1177/1524839904273387
- Schwärzel, L. (2016). *Prozessevaluation eines Projektes zum betrieblichen Gesundheitsmanagement*. In Wieland, R., Seiler, K. & Hammes, M. (Hrsg.), *Psychologie der Arbeitssicherheit und Gesundheit. Dialog statt Monolog*. (S.75-78). Krönig: Asanger Verlag.
- Stufflebeam, D. L. & Shinkfield, A. J. (2007). *Evaluation theory, models and applications*. San Francisco: Jossey-Bass.
- Stufflebeam, D. L. (2003a). *The CIPP Model for evaluation*. Presented at the 2003 Annual Conference of the Oregon Program Evaluators Network (OPEN). Portland, OR.
- Stufflebeam, D. L. (2003b). *The CIPP Model for evaluation*. In T. Kellaghan, D. L. Stufflebeam & L. A. Wingate (Hrsg.), *International handbook of educational evaluation. Part one: Perspectives* (Kluwer international handbooks of education, Bd. 9). Dordrecht: Kluwer.

- Stufflebeam, D. L. (2007). CIPP Evaluation Model checklist. A tool for applying the CIPP Model to assess long-term enterprises (2nd ed.). Verfügbar unter https://www.wmich.edu/sites/default/files/attachments/u350/2014/cippchecklist_mar07.pdf
- Stufflebeam, D.L. & Shinkfield, A.J. (2007). Evaluation Theory, Models, & Applications. SanFrancisco: Wiley.
- Watzl, F. (2016). Programmevaluation? Wo steht das im Gesetz? In Wieland, R., Seiler, K. & Hammes, M. (Hrsg.), Psychologie der Arbeitssicherheit und Gesundheit. Dialog statt Monolog. (S.79-82). Krönig: Asanger Verlag.
- Westermann, R. (2002). Merkmale und Varianten von Evaluationen: Überblick und Klassifikation. Zeitschrift für Psychologie, 210 (1), 4-26. doi: 10.1026//0044-3409.210.1.4
- Wheeler, D. und Sillanpää, M. (1997). The Stakeholder Corporation: A Blueprint for Maximizing Stakeholder Value. (London: Pitman)

7 Autor*innen

Monika Eigenstetter

Prof. Dr. Monika Eigenstetter ist seit 2009 Professorin für Arbeits- und Organisationspsychologie an der Hochschule Niederrhein, seit 2015 Leitung des Forschungsinstitut A.U.G.E. und seit 2013 des EthNa Kompetenzzentrum CSR. Ihre Arbeitsschwerpunkte Arbeitssicherheit und Gesundheit, Corporate Social Responsibility, Umweltschutz

Udo Keil

Dr. Udo Keil ist Akademischer Oberrat am Institut für Psychologie der Technischen Universität Darmstadt. Er ist Mitglied der Forschungsgruppe Arbeits- und Ingenieurpsychologie und Geschäftsführer des Instituts für Psychologie. Er lehrt an der TU-Darmstadt Arbeits- und Organisationspsychologie sowie Forschungs- und Evaluationsmethoden.

Herr Keil beschäftigt sich seit langer Zeit mit der Psychologie der Arbeitssicherheit und Gesundheit, in den letzten Jahren mit dem Fokus auf Gesundheitsmanagementansätze und deren Evaluation.

Werner Hamacher

Werner Hamacher (Dipl.-Päd. Mit Schwerpunkt Erwachsenenbildung) ist Geschäftsführer bei Systemkonzept – Gesellschaft für Systemforschung und Konzeptentwicklung mbH in Köln. Er ist seit 1980 im Arbeitsschutz tätig mit den Schwerpunkten Forschung, Beratung und Bildung insbesondere in der Konzeption und Durchführung der Sifa-Ausbildung, der Leitung und Durchführung von Forschungsprojekten (z. B. Sifa-Langzeitstudie), der Entwicklung der DGUV Vorschrift 2. Er befasst sich mit dem Thema Evaluation sowohl im Kontext von wissenschaftlichen Projekten als auch in der betrieblichen Praxis des Arbeitsschutzes sowie im Feld der Entwicklung und Wirksamkeit von Bildungsmaßnahmen und der Kompetenzentwicklung.

Tobias Löcher

Tobias Löcher (Dipl. Psychologe) arbeitet bei der Berufsgenossenschaft Rohstoffe und Chemische Industrie in der Erwachsenenbildung („Kompetenzcenter Qualifizierung“ am Standort Maikammer). Als Verhaltenswissenschaftler beschäftigt er sich übergeordnet mit der Steigerung des Wirkungsgrades von Seminaren. Inhaltlich engagiert er sich bzgl. klassischer Fragestellungen der sachförderlichen Verhaltensbeeinflussung im Sinne des Arbeitsschutzes. Das Thema Evaluation begleitet dabei das kontinuierliche Prüfen und Verbessern von Veranstaltungen und Herangehensweisen.

Lotte Schwärzel

Lotte Schwärzel (M.Sc. Psychologie) arbeitet bei der Berufsgenossenschaft Rohstoffe und chemische Industrie in der Prävention im Kompetenz-Center Gesundheit im Betrieb. Zu ihren Arbeitsschwerpunkten gehören Arbeitssicherheit und Gesundheit, gesundheitsförderliche Führung, Beschäftigungsfähigkeit sowie Evaluation. In ihrer vorherigen Tätigkeit an der Technischen Universität Darmstadt befasste sich Frau Schwärzel mit Gesundheitsmanagement und dessen Evaluation.

Frank Watzl

Frank Watzl berät als Fachkraft für Arbeitssicherheit Unternehmer und Beschäftigte zu Fragen der Sicherheit und des Gesundheitsschutzes. Er ist Referent in Seminaren für Arbeitsschutzakteure bei verschiedenen Bildungsträgern.

Zitationsvorschlag:

Hamacher Werner, Watzl Frank, Eigenstetter Monika, Schwärzel Lotte, Keil Udo, Löcher Tobias (2017). Evaluation oder „Tun wir die richtigen Dinge? – Tun wir die Dinge richtig?“ Köln.