



VON DER INVESTITIONS- ENTSCHEIDUNG ZUR WIRKSAMKEITSLOGIK

Investitionen begründen ist der erste Schritt. Wirkung nachweisen der entscheidende.

Dieses Whitepaper zeigt, wie Unternehmen aus Annahmen belastbare Wirksamkeitslogiken entwickeln und Gesunde Arbeit systematisch steuern können.

AUTOREN

Tim Klingmann & Alexander Maas

Executive Summary

Viele Unternehmen können heute plausibel begründen, warum sich Investitionen in Gesunde Arbeit lohnen. Deutlich schwieriger ist jedoch die Frage, welche Wirkung diese Investitionen tatsächlich entfalten und wie sich ihr Beitrag zu Unternehmenszielen nachvollziehbar belegen lässt. Genau an diesem Punkt setzt dieses Whitepaper an. Es beschreibt den Übergang vom Entscheidungs-Business-Case zum Steuerungs- und Nachweis-Business-Case und zeigt, wie aus Wirkungsannahmen eine belastbare Wirksamkeitslogik entwickelt werden kann.

Ausgangspunkt ist die Erkenntnis, dass ein perfekter, monokausaler Wirkungsnachweis im betrieblichen Kontext meist weder realistisch noch erforderlich ist. Faktoren wie Fehlzeiten, Produktivität, Mitarbeiterbindung oder Fluktuation werden von zahlreichen Einflüssen bestimmt und lassen sich selten eindeutig auf einzelne Maßnahmen zurückführen. Statt einer reinen Beweislogik verfolgt das Whitepaper deshalb eine Beitragslogik: Nicht die Frage, ob eine Maßnahme allein eine Wirkung verursacht hat, steht im Mittelpunkt, sondern ob sie nachweisbar zu einer positiven Entwicklung beigetragen hat. Dies orientiert sich am etablierten Ansatz der Contribution Analysis.

Als konzeptionelle Grundlage dient das MOOVE Value Driven Framework, das die Zusammenhänge zwischen Arbeitsbedingungen, Treiberfaktoren Gesunder Arbeit, subjektivem Arbeitserleben und organisationalen Value-Faktoren strukturiert. Dadurch wird nachvollziehbar, über welche Wirkpfade Investitionen in Gesunde Arbeit langfristig zu relevanten Unternehmenszielen wie Produktivität, Mitarbeiterbindung, Leistungsfähigkeit oder reduziertem Absentismus beitragen können.

Im Zentrum des Whitepapers steht eine praxisorientierte Fünf-Ebenen-Prüflogik, die Unternehmen bei der systematischen Wirksamkeitssteuerung unterstützt. Sie betrachtet nacheinander die Wirksamkeitsvoraussetzungen, die Qualität der Umsetzung, die Veränderung relevanter Treiber und Wirkmechanismen, die Entwicklung von Outcomes und Value-Faktoren sowie die Frage der Zuschreibung und ökonomischen Übersetzung. Durch diese Struktur wird sichtbar, ob eine Maßnahme tatsächlich die gewünschte Wirkungskette in Bewegung setzt oder ob Anpassungen notwendig sind.

Ein wesentlicher Mehrwert der Prüflogik besteht darin, dass Wirkung nicht erst anhand spät reagierender Kennzahlen wie Fehlzeiten oder Fluktuation bewertet wird. Bereits Veränderungen im Führungsverhalten, in der Arbeitsgestaltung, im subjektiven Wohlbefinden oder in anderen relevanten Treibern liefern frühzeitig Hinweise darauf, ob die eingeschlagene Intervention wirksam ist. Dadurch entsteht eine deutlich bessere Grundlage für Steuerungsentscheidungen als durch reine Teilnahme- oder Zufriedenheitswerte.

Die eigentliche Zielsetzung der Wirksamkeitsmessung besteht jedoch nicht in der Datenerhebung, sondern in einer besseren Entscheidungsfindung. Daten sollen helfen zu beurteilen, welche Maßnahmen fortgeführt, angepasst, ausgeweitet oder beendet werden sollten. Wirksamkeitssteuerung wird damit zu einem kontinuierlichen Lernprozess, der Investitionen gezielter macht, Ressourcen effizienter einsetzt und die Argumentation gegenüber Geschäftsführung und Stakeholdern stärkt.

Das Whitepaper zeigt zudem, dass der Einstieg in eine wirksamkeitsorientierte Steuerung nicht komplex sein muss. Bereits vorhandene Daten, eine klar definierte Zielgruppe, eine nachvollziehbare Wirkungshypothese und wenige relevante Kennzahlen reichen aus, um erste belastbare Erkenntnisse zu gewinnen. Mit zunehmender Datenreife kann die Wirksamkeitssteuerung schrittweise erweitert werden.

Die zentrale Botschaft lautet: Der Erfolg Gesunder Arbeit zeigt sich nicht in der Anzahl umgesetzter Maßnahmen, sondern in ihrem nachvollziehbaren Beitrag zu besseren Arbeitsbedingungen, einem positiven Arbeitserleben und relevanten Unternehmenszielen. Unternehmen, die Wirkung entlang einer systematischen Wirksamkeitslogik steuern, schaffen die Grundlage für fundiertere Investitionsentscheidungen, höhere Transparenz und eine nachhaltigere Wertschöpfung durch Gesunde Arbeit.

Inhaltsverzeichnis

Von der Investitionsentscheidung zur Wirksamkeitslogik	S. 4
Rückblick – der Business Case in 2 Stufen	S. 4
Das zentrale Problem	S. 5
Zielstellung/Positionierung dieses Whitepapers	S. 6
Das MOOVE Value Driven Framework als Basis für die	
Wirksamkeitslogik	S. 6
Die Fünf-Ebenen Prüflogik	S. 7
Ein Praxisfall, der alles verbindet	S. 8
Ebene 1 – Wirksamkeitsvoraussetzungen	S. 9
Ebene 2 – Umsetzungsqualität	S. 11
Ebene 3 – Treiber und Wirkmechanismen	S. 12
Ebene 4 – Value-Faktoren und Outcome	S. 16
Ebene 5 – Zuschreibung und ökonomische Übersetzung	S. 17
Von Daten zu Entscheidungen	S. 19
Pragmatische Einstiege	S. 20
Ausblick – Vom Kennzahlenset zum Wirksamkeitscockpit	S. 23
Fazit	S. 24
Quellenverzeichnis	S. 25

Von der Investitionsentscheidung zur Wirksamkeitslogik

Das Whitepaper „Business Case Gesunde Arbeit“ zeigt, wie Unternehmen Investitionen in Gesunde Arbeit auch ohne vollständige Kennzahlen plausibel begründen können. Im Mittelpunkt stehen wissenschaftlich fundierte Wirkungsannahmen, realistische Bandbreiten und Szenarien, die eine Investitionsentscheidung auch unter Unsicherheit ermöglichen. Der dort beschriebene Entscheidungs-Business-Case beantwortet damit vor allem die Frage, ob und in welcher Größenordnung in Gesunde Arbeit investiert werden sollte.

Mit der damit verbundenen Budgetfreigabe verändert sich jedoch die Perspektive. Ab diesem Punkt geht es nicht mehr nur darum, eine Investition plausibel zu begründen, sondern ihre Wirkung im Zeitverlauf nachvollziehbar zu machen. Denn früher oder später entsteht die Frage, was die Investition tatsächlich gebracht hat, ob relevante Zielgruppen erreicht wurden, ob sich Belastungen, Fehlzeiten oder Fluktuation verändern oder ganz pauschal, ob Gesunde Arbeit im Unternehmen wirksam wird.

Genau hier setzt der zweite Schritt an. Aus dem Entscheidungs-Business-Case wird ein Steuerungs- und Nachweis-Business-Case, der nicht mehr primär auf Annahmen und Szenarien basiert, sondern die getätigten Annahmen mit echten Unternehmensdaten überprüft, schärft und für weitere Entscheidungen nutzbar macht.

Der Business Case Gesunde Arbeit entfaltet seine Wirkung damit in zwei Stufen. In der ersten Stufe geht es um die Investitionsentscheidung auf Basis plausibler Szenarien, Bandbreiten und Wirkungsannahmen. In der zweiten Stufe geht es darum, Wirkung sichtbar zu machen, Daten systematisch einzuordnen und aus der ursprünglichen Investitionslogik eine belastbare Wirksamkeitslogik zu entwickeln.

Der Business Case in 2 Stufen

Der Business Case Gesunde Arbeit entfaltet seine Wirkung in zwei Stufen. Dieses Whitepaper fokussiert sich auf Stufe 2.

Stufe 1

Entscheidungs-Business-Case: Szenarien, Bandbreiten, Wirkungsannahmen → Investitionsentscheidung

Stufe 2

Steuerungs- & Nachweis-Business-Case: Wirkung sichtbar machen, Annahmen mit echten Daten schärfen

Das zentrale Problem

Viele HR- und BGM-Teams sind hinsichtlich der zweiten Stufe des Business Case heute nur wenig aussagefähig. Häufig fehlen zunächst grundlegende Voraussetzungen wie klare Ziele, passende Kennzahlen, einfache Wirkungsketten und eine nachvollziehbare Verbindung zwischen Maßnahmen, Arbeitserleben und Unternehmenskennzahlen.

Dabei sollte diese begrenzte Aussagefähigkeit vieler HR- und BGM-Teams nicht vorschnell als methodische Schwäche verstanden werden. Der Wirksamkeitsnachweis bleibt auch dann anspruchsvoll, wenn Gesunde Arbeit bereits professionell und systematisch aufgestellt ist, weil Unternehmen sich in einem komplexen, dynamischen Umfeld mit vielen Nachweisbarrieren befinden.

Diese Nachweisbarrieren ergeben sich vor allem aus der Komplexität der zugrunde liegenden Wirkzusammenhänge. Value-Faktoren wie Fehlzeiten, Fluktuation, Produktivität oder Bindung sind multikausal bestimmt, Wirkungen entfalten sich häufig über längere Zeiträume, Teilnehmende unterscheiden sich nicht selten systematisch von Nicht-Teilnehmenden, zentrale Konstrukte wie Belastung, Wohlbefinden oder Führungsqualität sind methodisch anspruchsvoll zu messen und ohne Vergleich bleibt offen, was auch ohne Maßnahme passiert wäre.

In der Praxis führen diese Bedingungen und Barrieren dazu, dass ein perfekter, monokausaler Wirkungsnachweis im betrieblichen Kontext meist weder realistisch noch erforderlich ist. Vielmehr braucht es einen anderen Anspruch.



Der wissenschaftlich saubere Ausweg ist nicht der Verzicht auf den Nachweis, sondern der Wechsel von der Beweis- zur Beitragslogik.

Statt vorschnell zu behaupten, eine Maßnahme habe eine Unternehmenskennzahl allein verursacht, muss gezeigt werden, ob eine Investition plausibel und nachvollziehbar zu einer Veränderung beigetragen hat. Dafür muss die zugrunde gelegte Wirkungskette empirisch gestützt, mit mehreren Datenpunkten geprüft und gegenüber konkurrierenden Erklärungen eingeordnet werden. Diese beitragsorientierte Perspektive ist als Contribution Analysis methodisch etabliert und bildet den roten Faden dieses Whitepapers (Mayne, 2001).

Zielstellung/ Positionierung dieses Whitepapers

Aufgrund der beschriebenen Komplexität legt dieses Whitepaper den Fokus auf eine grundlegende Wirksamkeitslogik, die HR- und BGM-Teams dabei unterstützt, Investitionen in Gesunde Arbeit fundierter intern zu begründen, passende Evaluationsstrukturen aufzubauen und erste Schritte in Richtung einer belastbaren Wirksamkeitssteuerung zu gehen.

Da jedes Unternehmen mit eigenen Rahmenbedingungen, Datenlagen und Fragestellungen arbeitet, ersetzt dieses Whitepaper keine individuelle Analyse.

ZIELGRUPPE

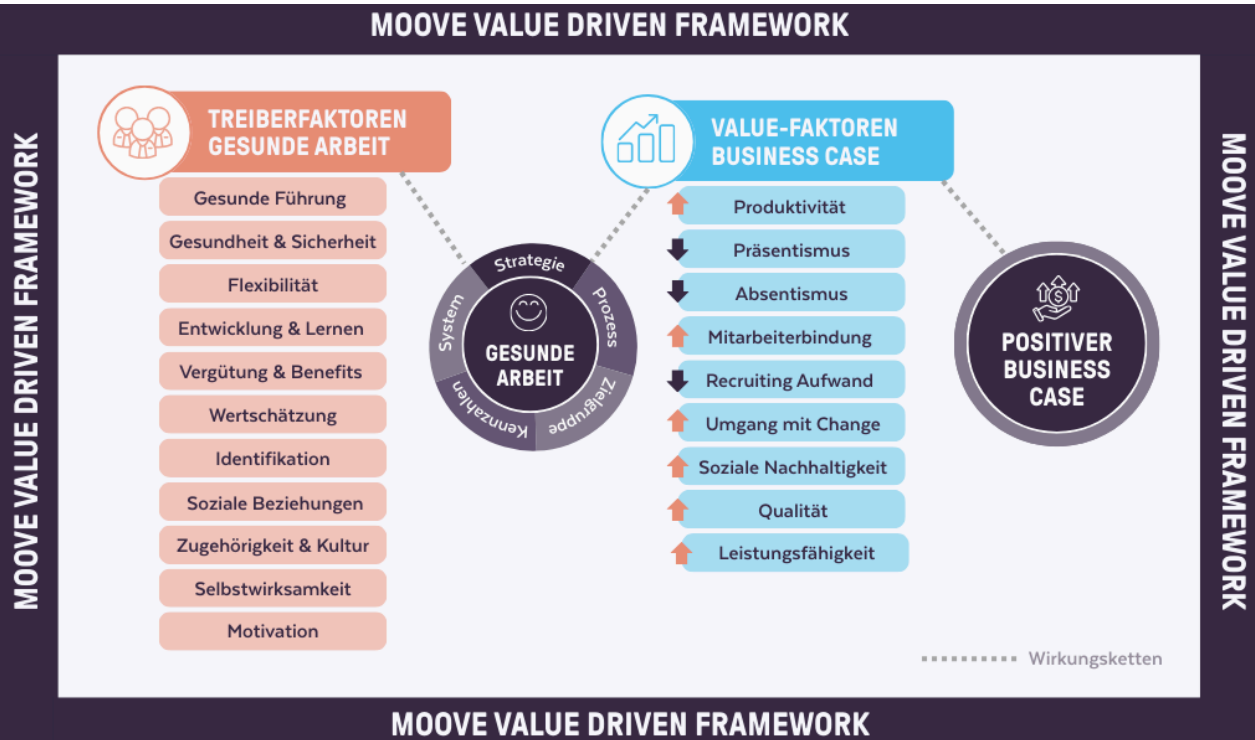
Adressiert sind HR- und BGM-Verantwortliche, die eine Investition bereits begründet haben und nun Wirkung nachvollziehbar machen wollen, ebenso Führungskräfte und Steuerungsgremien, die Ergebnisse einordnen müssen. Wer schnell in die Umsetzung kommen will, beginnt bei den pragmatischen Einstiegen, während die Fünf-Ebenen-Prüflogik den methodischen Kern des Whitepapers beschreibt.

Das MOOVE Value Driven Framework als Basis für die Wirksamkeitslogik

Für eine beitragsorientierte Wirksamkeitssteuerung braucht es zunächst ein Erklärmodell, das komplexe Wirkzusammenhänge strukturiert und für die betriebliche Praxis handhabbar macht. Das MOOVE Value Driven Framework bildet dafür den übergeordneten Bezugsrahmen. Es verbindet Erkenntnisse aus Wellbeing-Forschung, Arbeitspsychologie und theoriegeleiteten Wirkmodellen mit einer anwendungsorientierten Steuerungslogik.

Ausgangspunkt ist die Annahme, dass Wirkungen Gesunder Arbeit über mehrere, miteinander verbundene Einflussgrößen entstehen. Arbeitsbedingungen und zentrale Treiber wie beispielsweise Gesunde Führung, Gesundheit & Sicherheit, soziale Beziehungen oder Selbstwirksamkeit beeinflussen, wie Beschäftigte ihre Arbeit erleben und wie sich Gesundheit und Wohlbefinden entwickeln. Diese stehen wiederum mit organisationalen Value-Faktoren in Zusammenhang, über die sich der mögliche Beitrag Gesunder Arbeit zu relevanten Unternehmens- und Personalzielen abbilden lässt, etwa Produktivität, Absentismus, Präsentismus oder Mitarbeiterbindung in Zusammenhang (Bakker & Demerouti, 2007; De Neve & Ward, 2025; Niks et al., 2022).

In Anlehnung an das wissenschaftliche Konzept der Theory of Change (vgl. Rogers, 2014) lassen sich daraus plausible Wirkungspfade formulieren: Welche Veränderung soll erreicht werden, welche Treiber müssen sich dafür verändern und über welche Zwischenschritte kann daraus ein Beitrag zu organisationalen Ergebnissen entstehen? Das Framework schafft damit die konzeptionelle Grundlage, um Wirkungsannahmen explizit zu machen und anschließend systematisch zu prüfen.



Die Fünf-Ebenen Prüflogik

Das Value Driven Framework beschreibt den übergeordneten Wirkungszusammenhang. Für die praktische Anwendung braucht es darauf aufbauend eine Prüflogik, die sichtbar macht, ob konkrete Maßnahmen, Programme oder Lösungsarchitekturen einer plausiblen Wirkungskette folgen, wirksam umgesetzt werden und sich in relevanten Daten nachvollziehen lassen.

Die Fünf-Ebenen-Prüflogik strukturiert diesen Prozess schrittweise. Sie beginnt bei den Voraussetzungen für Wirksamkeit, betrachtet anschließend Umsetzung und erreichte Veränderungen bei relevanten Treibern und führt über weiterführende Outcomes bis zu organisationalen Value-Faktoren. Ergänzend wird geprüft, wie plausibel beobachtete Veränderungen der Intervention zugerechnet werden können und welche Konsequenzen sich daraus für Steuerung, Anpassung oder Weiterentwicklung ableiten lassen.

WIRKSAMKEITSVORAUSSETZUNGEN

UMSETZUNGSQUALITÄT

TREIBER/ WIRKMECHANISMEN

VALUE-FAKTOREN/ OUTCOME

ZUSCHREIBUNG & ÖKONOMISCHE ÜBERSETZUNG

Ein Praxisfall, der alles verbindet

Damit die Fünf-Ebenen-Prüflogik greifbar wird, wird sie im Folgenden anhand eines konkreten, bewusst vereinfachten Fallbeispiels durchgängig angewendet. So wird sichtbar, wie sich die einzelnen Ebenen in der Praxis aufeinander beziehen und für die Wirksamkeitsbewertung nutzen lassen.



Der Fall: Führungskräfteprogramm im Produktionsbereich

Ausgangslage:

In einem belasteten Produktionsbereich (120 Mitarbeitende, 15 Führungskräfte) zeigen sich erhöhte Kurzzeit-Fehlzeiten, kritische Werte in der Gefährdungsbeurteilung, hohe Arbeitsintensität, unklare Prioritäten und erste Hinweise auf sinkende Bleibeabsicht.

Naheliegende, aber vorschnelle Reaktion: „Wir buchen ein Stress- oder Resilienzseminar.“

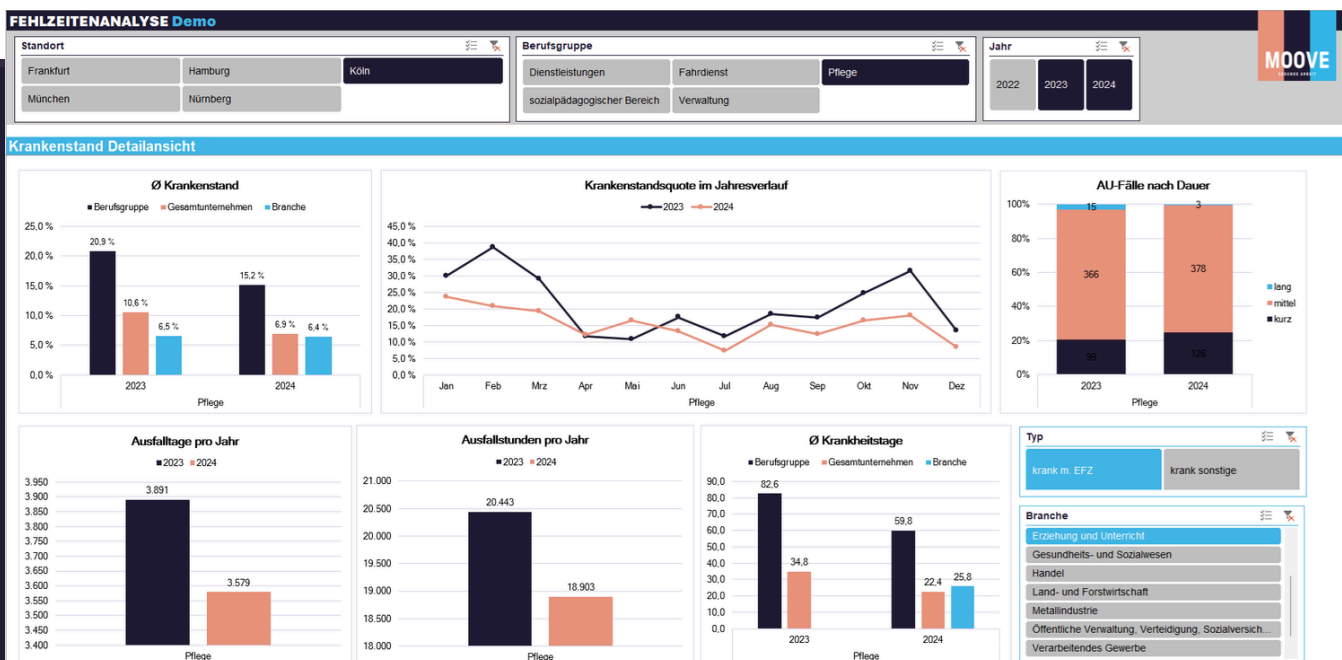
Bessere Entscheidung:

Zunächst klären, welcher Hebel wirklich verändert werden soll – hier Führung, Priorisierung und Rollenklarheit. Daraus wird ein Führungskräfteprogramm (Priorisierung, unterstützende Kommunikation, regelmäßige Belastungsgespräche).

Wirksamkeitsvoraussetzungen

Auf der ersten Ebene wird geprüft, ob eine Maßnahme überhaupt gut genug vorbereitet ist, um später sinnvoll über Wirkung sprechen zu können. Viele Vorhaben im Kontext Gesunder Arbeit starten jedoch vorschnell mit einer Lösung, etwa mit einem Stressseminar, einem Führungskräfte-Training oder einem digitalen Angebot, bevor klar ist, welches konkrete Problem damit bearbeitet werden soll.

Für die Wirksamkeitssteuerung ist deshalb der erste Schritt, die Ausgangslage zu klären. Es muss erkennbar sein, wo Handlungsbedarf liegt, welche Zielgruppe betroffen ist und welcher Hebel verändert werden soll. Hinweise können zum Beispiel aus der psychischen Gefährdungsbeurteilung, Fehlzeitenanalysen, Mitarbeitendenbefragungen, BEM-Auswertungen, Führungskräftefeedback oder Workshops kommen.



Für die Klärung der Ausgangslage können Fehlzeiten-Dashboards oder vertiefende Fehlzeitenanalysen wichtige Orientierung geben. Sie zeigen, in welchen Bereichen, Zeiträumen oder Beschäftigtengruppen Auffälligkeiten bestehen und helfen damit, Handlungsbedarfe präziser einzugrenzen. Für die Wirksamkeitslogik sind sie jedoch zunächst Ausgangsdaten und noch kein Nachweis für Ursache oder Wirkung.

Wenn der relevante Hebel und der fachliche Schwerpunkt definiert sind, folgt die nächste zentrale Frage

“Wie erklären wir uns die Wirkung?”

Denn auch eine gut gewählte Maßnahme wird erst dann steuerbar, wenn nachvollziehbar ist, über welche Zwischenschritte sie Veränderung im Arbeitsalltag auslösen soll.

An dieser Stelle wird die Wirkungshypothese formuliert. Dabei wird geprüft, ob Problem, Zielgruppe, Zielsetzung und Wirkungshypothese klar beschrieben sind und ob nachvollziehbar wird, welches Problem bearbeitet werden soll, welche Zielgruppe relevant ist, welcher Treiber sich verändern soll, welcher Value-Faktor mittelbar betroffen ist und in welchem Zeitraum erste Veränderungen realistisch sind.

Grundlage ist idealerweise eine Wirkungskette, die die Hypothese als nachvollziehbare Abfolge von Zwischenschritten sichtbar macht.



Wirkungskette & Theory of Change

Eine Wirkungskette stellt in wenigen plausiblen Schritten dar, was eine Maßnahme unmittelbar verändert, was daraus folgt und wo sich am Ende eine Veränderung im Arbeitserleben oder in einer Value-nahen Größe zeigt. Ihr Wert liegt darin, dass jeder Schritt später daraufhin überprüfbar ist, ob er tatsächlich eingetreten ist.

Der Gedanke entspricht der Theory of Change aus der Evaluationsforschung: Bei komplexen Interventionen sollten nicht nur Endergebnisse betrachtet werden, sondern auch Implementierung, Kontext und Wirkmechanismen (Moore et al., 2015; Nielsen & Randall, 2013) – hier bewusst in schlanker, betrieblich handhabbarer Form.

Wirkungskette am Fallbeispiel



Die Wirkungskette verdeutlicht, warum das Programm nicht unmittelbar an der Fluktuationsquote gemessen werden sollte. Plausibler ist ein mehrstufiger Wirkungsverlauf, bei dem sich zunächst das Führungsverhalten im Arbeitsalltag verändert und Mitarbeitende dadurch mehr Unterstützung, Klarheit und psychologische Sicherheit erleben. Werden Belastungen auf dieser Grundlage früher sichtbar und besser bearbeitbar, kann sich das Arbeitserleben schrittweise verbessern. Eine stabilere Bleibeabsicht oder ein geringeres Fluktuationsrisiko wären daher eher als spätere Hinweise auf Wirkung zu interpretieren.

QUICK-TIPP

Einheitliche Filterung

Kennzahlen werden nur vergleichbar, wenn Zeitraum, Zielgruppe und Organisationseinheit konsistent abgebildet werden. Filtern Sie alle Datenquellen (Gefährdungsbeurteilung, Befragung, Fehlzeiten-Dashboard, Maßnahmenfragebogen) nach denselben Kategorien, z. B. Jahr/Quartal, Bereich, Standort, Funktion, Zielgruppe. Nur so lassen sich einzelne Datenpunkte später logisch verbinden.

EBENE 2

Umsetzungsqualität

Auf der zweiten Ebene wird geprüft, ob eine Maßnahme so umgesetzt wurde, dass Wirkung überhaupt entstehen konnte. Das klingt banal, ist in der Praxis aber einer der wichtigsten Punkte der Wirksamkeitssteuerung.

Am Fallbeispiel

Nach sechs Monaten zeigen sich kaum Veränderungen in der Befragung. Bevor das Programm als „unwirksam“ gilt, prüfen wir: Haben genug Führungskräfte aus dem Bereich teilgenommen? Wurden Transferaufgaben umgesetzt? Trug die Bereichsleitung die Umsetzung mit? Hatte das Programm Bezug zu den realen Belastungssituationen?

Der Kern dieser Ebene ist die Unterscheidung zwischen einem Wirkungsproblem und einem Umsetzungsproblem:

- Wirkungsproblem: Die Zielgruppe wurde erreicht, die Intervention wie vorgesehen umgesetzt und die angenommenen Wirkmechanismen wurden aktiviert – dennoch zeigt sich keine relevante Veränderung. In diesem Fall sollte die zugrunde liegende Wirkungshypothese und damit die fachliche Logik der Intervention überprüft werden.
- Umsetzungsproblem: Die Zielgruppe wurde nur unzureichend erreicht, der Transfer in den Arbeitsalltag gelingt nicht oder relevante Akteure wie Führungskräfte tragen die Umsetzung nicht ausreichend mit. In diesem Fall ist nicht zwingend die Intervention selbst infrage zu stellen, sondern vor allem ihre Umsetzung und organisationale Verankerung zu verbessern.

Anders ausdrückt: Ohne Ebene 2 wird ein schlecht umgesetztes Programm mit einem fachlich falschen verwechselt. Damit besteht die Gefahr, eine grundsätzlich passende Intervention zu verwerfen, obwohl nicht ihre fachliche Logik, sondern ihre Umsetzung das eigentliche Problem ist.

4 Fragen helfen in der Praxis

1. Wurde die relevante Zielgruppe in ausreichender Breite und Tiefe erreicht?
2. Wurde die Maßnahme vollständig, qualitätsgesichert und wie geplant umgesetzt?
3. Wurde die Umsetzung durch Führung, HR, BGM und Organisation sichtbar unterstützt?
4. Gab es konkrete Transferimpulse, die eine Verankerung im Arbeitsalltag ermöglicht haben?

Wissenschaftlicher Einwurf: Prozessevaluation

Diese Fragen knüpfen an etablierte Ansätze der Prozessevaluation an. Das RE-AIM-Rahmenwerk unterscheidet beispielsweise Reichweite, Wirksamkeit, Übernahme, Umsetzung und Verstetigung. Die MRC-Leitlinie zur Prozessevaluation komplexer Interventionen betont ergänzend, dass neben Ergebnissen auch Reichweite, Dosis, Umsetzungstreue, Kontext und aktivierte Wirkmechanismen betrachtet werden sollten (Glasgow, Vogt & Boles, 1999; Moore et al., 2015).

QUICK-TIPP

Standardisierte Maßnahmenevaluation

Bewerten Sie Maßnahmen nicht jedes Mal anders – sonst entsteht kein kumulatives Wissen. Ein einheitlicher Kurzfragebogen (Relevanz, Zielgruppenpassung, Anwendbarkeit, Transfer, erste wahrgenommene Wirkung) nach jeder größeren Maßnahme schafft Vergleichbarkeit und eine Lernkurve über Formate hinweg.

Treiber und Wirkmechanismen

Auf der dritten Ebene wird geprüft, ob sich die Treiber und Wirkmechanismen verändern, über die eine Maßnahme ihre Wirkung entfalten sollen. Diese Ebene ist besonders wertvoll, da sie zwischen Umsetzung und späten Ergebniskennzahlen liegt und viel früher Hinweise liefert als Fehlzeiten oder Fluktuation.

Dabei ist zwischen Treibern und Wirkmechanismen zu unterscheiden.

Treiber sind die im MOOVE Value Driven Framework definierten Einflussbereiche Gesunder Arbeit. Dazu gehören Gesunde Führung, Gesundheit & Sicherheit, Flexibilität, Entwicklung & Lernen, Vergütung & Benefits, Wertschätzung, Identifikation, soziale Beziehungen, Zugehörigkeit & Kultur, Selbstwirksamkeit und Motivation. Sie bilden die übergreifende Wirklogik und schaffen einen einheitlichen Bezugsrahmen über unterschiedliche Interventionen hinweg.

Wirkmechanismen beschreiben dagegen, wie eine konkrete Maßnahme einen solchen Treiber verändern soll. Sie übersetzen die Intervention in konkrete Veränderungen des Arbeitsalltags. Bei einem Führungskräfteprogramm kann der Wirkmechanismus beispielsweise darin bestehen, dass Führungskräfte Prioritäten klarer kommunizieren, Belastungen früher ansprechen oder Mitarbeitende stärker in Problemlösungen einbeziehen. Diese Veränderungen sollen anschließend den Treiber Gesunde Führung verbessern.

Am Fallbeispiel

Beim Führungskräfteprogramm reicht es deshalb nicht, Teilnahme und Zufriedenheit zu erfassen. Zunächst ist zu prüfen, ob sich das angestrebte Verhalten im Arbeitsalltag verändert: Werden Prioritäten klarer besprochen? Werden Belastungen früher aufgegriffen? Erhalten Mitarbeitende mehr Orientierung und Unterstützung?

Darauf aufbauend wird untersucht, ob sich der adressierte Treiber Gesunde Führung aus Sicht der Mitarbeitenden verbessert. Je nach Wirkungshypothese können auch weitere Treiber wie Wertschätzung, Selbstwirksamkeit oder soziale Beziehungen relevant sein.

Erst wenn sich die adressierten Treiber verändern, wird plausibler, dass sich anschließend auch Arbeitserleben, Arbeitsfähigkeit und später organisationale Value-Faktoren verändern.

Für die praktische Steuerung muss eine Wirkungshypothese nicht in ihrer gesamten Breite gemessen werden. Entscheidend ist, pro Intervention die wenigen Treiber auszuwählen, die für den angenommenen Wirkungsweg tatsächlich zentral sind, und zu konkretisieren, woran sich ihre Veränderung im Arbeitsalltag erkennen lässt. In der Regel reichen dafür ein bis drei Treiber, wie beispielhaft in der nachfolgenden Tabelle dargestellt.

Maßnahme	Wirkmechanismus im Arbeitsalltag	Treiber	Mögliche Prüffrage
Führungskräfteprogramm	Führungskräfte klären Prioritäten, geben Orientierung und sprechen Belastungen frühzeitig an.	Gesunde Führung	„Meine Führungskraft unterstützt mich dabei, gesund und leistungsfähig zu arbeiten.“
Maßnahmen aus der psyGBU	Aufgaben, Kapazitäten, Unterbrechungen und Erholungsmöglichkeiten werden gesundheitsgerechter gestaltet.	Gesundheit & Sicherheit	„Meine Arbeitsbedingungen ermöglichen es mir, langfristig gesund zu arbeiten.“
Flexible Arbeitsgestaltung	Mitarbeitende erhalten größere Spielräume bei Arbeitszeit und Arbeitsort und können Anforderungen besser mit ihrer Lebenssituation verbinden.	Flexibilität	„Die Gestaltung meiner Arbeit bietet mir ausreichend Flexibilität.“
Team-Workshop	Zusammenarbeit, gegenseitige Unterstützung und offene Kommunikation werden im Team gestärkt.	Soziale Beziehungen	„In meinem Arbeitsumfeld unterstützen wir uns gegenseitig.“
Feedback- und Anerkennungsformat	Beiträge und Leistungen werden regelmäßiger wahrgenommen, rückgemeldet und anerkannt.	Wertschätzung	„Meine Leistung und mein Beitrag werden angemessen wertgeschätzt.“

Durch die Orientierung an der Wirkungshypothese wird erkennbar, ob eine Intervention lediglich umgesetzt wurde oder ob sie tatsächlich die vorgesehenen Veränderungen im Arbeitsalltag und bei den adressierten Treibern anstößt. Das schafft eine wesentlich belastbarere Steuerungsgrundlage als Teilnahme- oder Zufriedenheitswerte allein. Bleibt die erwartete Veränderung aus, lässt sich zudem früher prüfen, ob der Wirkmechanismus nicht greift, die Intervention anders ausgestaltet werden muss oder die ursprüngliche Wirkungshypothese zu überdenken ist.



Subjektives Wohlbefinden als Brücke zu Ebene 4

Neben den adressierten Treibern können sich erste Veränderungen auch im subjektiven Arbeitserleben zeigen. Solche näheren Outcomes bilden die Brücke zu Ebene 4 und reagieren häufig früher als organisationale Value-Faktoren wie Absentismus, Mitarbeiterbindung oder Produktivität. Subjektives Wohlbefinden kann deshalb eine wichtige Brückenkenzahl innerhalb der Wirkungskette sein:

VDF-Treiber → subjektives Wohlbefinden und Arbeitserleben → Value-Faktoren

Die Forschung unterscheidet dabei verschiedene Facetten des subjektiven Wohlbefindens. Für den Arbeitskontext sind insbesondere drei Perspektiven relevant:

Drei Dimensionen (Oxford Wellbeing Research Centre):

1. Bewertend: Wie beurteilen Beschäftigte ihre Arbeit insgesamt, beispielsweise hinsichtlich ihrer Arbeitszufriedenheit?
2. Affektiv: Wie erleben Beschäftigte ihre Arbeit emotional, beispielsweise mit Blick auf positive Gefühle oder Stress?
3. Eudaimonisch: In welchem Maß wird die eigene Arbeit als sinnvoll und bedeutsam erlebt?

Übersetzt in konkrete Items empfiehlt die englische Forschungsgruppe:

Job Satisfaction: „I feel completely satisfied with my work“

Happiness: “I feel happy at work, most of the time”

Stress: “I feel stressed at work, most of the time”

Purpose” “My work has a clear sense of purpose”

Die vier Items sind ein bewusst schlanker Einstieg und beispielhaft zu verstehen. Je nach Fragestellung können ergänzend validierte Instrumente wie WHO-5, COPSOQ oder UWES eingesetzt werden. Dabei sollte das gewählte Instrument immer zu dem Konstrukt passen, das tatsächlich untersucht werden soll.

Knüpft an die Subjective-Wellbeing-Forschung an (Diener, 1984; De Neve & Ward, 2023).

Value-Faktoren und Outcome

Auf der vierten Ebene wird geprüft, ob sich die für die Wirkungshypothese relevanten Outcomes und Value-Faktoren in die erwartete Richtung entwickeln. Im Whitepaper Business Case Gesunde Arbeit (Klingmann, 2025) wurden ausgewählte Faktoren wie Absentismus, Präsentismus, Fluktuation, Produktivität und Recruiting genutzt, um wirtschaftliche Größenordnungen und potenzielle Nutzenbeiträge plausibel abzuschätzen. Im Steuerungs- und Nachweis-Business-Case verändert sich die Perspektive: Nun wird beobachtet, ob sich entsprechende Ergebniskennzahlen im Zeitverlauf tatsächlich verändern.

Dabei können zunächst nähere Outcomes wie Wohlbefinden, Arbeitsfähigkeit, Erschöpfung oder Bleibeabsicht relevant sein. Daran anschließend richtet sich der Blick auf die organisationalen Value-Faktoren des MOOVE Value Driven Framework, beispielsweise Produktivität, Präsentismus, Absentismus, Mitarbeiterbindung, Recruiting-Aufwand, Qualität oder Leistungsfähigkeit.

Diese Ergebnisgrößen liegen weiter hinten in der Wirkungskette als Umsetzung, Wirkmechanismen und Treiber. Veränderungen treten deshalb häufig zeitverzögert auf und werden zugleich von zahlreichen weiteren Faktoren beeinflusst, etwa durch Saisonalität, Altersstruktur, Arbeitsmarkt, Auftragslage oder organisationale Veränderungen.

Am Fallbeispiel

Beim Führungskräfteprogramm wurde angenommen, dass Führungskräfte Prioritäten klarer kommunizieren, Belastungen früher ansprechen und Mitarbeitende gezielter unterstützen. Wenn sich dadurch der Treiber Gesunde Führung verbessert, sollte sich dies zunächst auch im Erleben der betroffenen Mitarbeitenden zeigen – beispielsweise durch geringere mentale Erschöpfung und eine stabilere Bleibeabsicht.

Erst mit größerem zeitlichem Abstand wird geprüft, ob sich dieser Wirkungsweg auch in den relevanten Value-Faktoren fortsetzt. Für das Führungskräfteprogramm könnten dies insbesondere Präsentismus, Mitarbeiterbindung und Absentismus sein. Recruiting-Aufwand wäre erst dann relevant, wenn sich eine verbesserte Mitarbeiterbindung tatsächlich in weniger Abgängen und damit geringerem Ersatzbedarf niederschlägt.

Für die Auswertung sind zwei Prinzipien zentral. Erstens werden nur solche Outcomes und Value-Faktoren betrachtet, die aus der zuvor formulierten Wirkungshypothese ableitbar sind. Eine nachträglich ausgewählte Kennzahl, die sich zufällig positiv entwickelt, liefert keinen belastbaren Wirkungsnachweis. Zweitens muss die Auswertung auf der Ebene erfolgen, auf der eine Wirkung plausibel zu erwarten ist. Richtet sich eine Intervention beispielsweise an einen bestimmten Bereich oder eine bestimmte Zielgruppe, sollte auch dort ausgewertet werden. Im Unternehmensdurchschnitt können relevante Veränderungen sonst leicht verschwinden.



Grundaussage

Ebene 4 zeigt, ob sich die erwartete Entwicklung über die Wirkungskette bis zu den relevanten Outcomes und Value-Faktoren fortsetzt. Ein Beleg für Kausalität ist damit noch nicht erbracht. Zeigen Umsetzung, Wirkmechanismen, Treiber und Ergebnisse jedoch konsistent in dieselbe Richtung, verdichtet sich die Evidenz für einen plausiblen Wirkungsbeitrag. Die Frage, wie belastbar dieser Beitrag der Intervention zugeschrieben werden kann, folgt auf der nächsten Ebene.

EBENE 5

Zuschreibung und ökonomische Übersetzung

Auf der fünften Ebene geht es um die anspruchsvollste Frage: Wie plausibel lässt sich die beobachtete Veränderung der Intervention zurechnen und welcher wirtschaftliche Beitrag lässt sich daraus ableiten?

Die Zuschreibung beginnt dabei nicht erst auf dieser Ebene. Je konsistenter sich die erwartete Entwicklung bereits über Umsetzung, Wirkmechanismen, Treiber, Outcomes und Value-Faktoren hinweg zeigt, desto plausibler wird ein Wirkungsbeitrag der Intervention. Ebene 5 ergänzt diese Wirkungskette um methodische Prüfungen, mit denen alternative Erklärungen systematisch eingeordnet werden können.

Dabei folgt der Ansatz einer Beitragslogik: Ziel ist nicht, unter Alltagsbedingungen einen zweifelsfreien Kausalitätsbeweis zu führen, sondern die Stärke der Evidenz für einen Wirkungsbeitrag nachvollziehbar zu erhöhen.

Drei methodische Ansätze sind dafür besonders praktikabel:

#1

Gestaffelter Rollout

Bereiche oder Teams starten zu unterschiedlichen Zeitpunkten. Zeigen sich Veränderungen jeweils nach Beginn der Intervention, während noch nicht gestartete Gruppen zunächst stabil bleiben, stärkt dies die Zuschreibung.

#2

Vergleichsgruppe

Die Entwicklung der Interventionsgruppe wird einer möglichst vergleichbaren Gruppe gegenübergestellt. Entscheidend ist dabei nicht nur, ob sich die Interventionsgruppe verbessert, sondern ob sich ihre Entwicklung erkennbar von der Vergleichsgruppe unterscheidet.

#3

Trendverlauf

Mehrere Messpunkte vor und nach der Intervention zeigen, ob sich ein bestehender Verlauf nach Beginn der Intervention verändert. Ein zeitlich passender Trendbruch kann die Wirkungshypothese zusätzlich stützen.

Ergänzend werden alternative Erklärungen systematisch geprüft, beispielsweise Saisonalität, Veränderungen der Personalstruktur, Reorganisationen, Auftragslage oder parallellaufende Initiativen. Je besser solche Einflussfaktoren eingeordnet werden können, desto belastbarer lässt sich der beobachtete Beitrag bewerten.

Am Fallbeispiel

Das Führungskräfteprogramm wird gestaffelt in drei Bereichen eingeführt. Im zuerst gestarteten Bereich verbessern sich nach Programmstart zunächst die auf Ebene 3 betrachteten Wirkmechanismen und der Treiber Gesunde Führung. Mit zeitlichem Abstand gehen auch mentale Erschöpfung und Präsentismus zurück, während ein noch nicht gestarteter, vergleichbarer Bereich zunächst weitgehend stabil bleibt.

Nach Einführung des Programms zeigt sich dort ein ähnliches Muster. Parallel werden Veränderungen der Arbeitslast, Personalwechsel und andere laufende Initiativen berücksichtigt. Das Zusammenspiel aus zeitlicher Abfolge, konsistenter Wirkungskette und Vergleich zwischen den Bereichen stärkt die Plausibilität, dass das Programm zu der beobachteten Entwicklung beigetragen hat.

Der ökonomische Übersetzungsschritt

Erst auf dieser Grundlage erfolgt die ökonomische Übersetzung. Sie knüpft an die im Value Driven Framework definierten Value-Faktoren an, beispielsweise Absentismus, Präsentismus, Mitarbeiterbindung, Produktivität oder Recruiting-Aufwand. Monetarisiert werden nur solche Veränderungen, für die ein wirtschaftlicher Zusammenhang und ein plausibler Wirkungsbeitrag begründet werden können.

Die Berechnung sollte konservativ erfolgen: Annahmen werden transparent gemacht, Unsicherheiten über Bandbreiten abgebildet und nur der plausibel zurechenbare Anteil berücksichtigt. So entsteht keine scheinexakte Renditekennzahl, sondern eine nachvollziehbare Größenordnung des wirtschaftlichen Beitrags. Diese kann anschließend in den nächsten Entscheidungs-Business-Case einfließen.

Von Daten zu Entscheidungen

Die beschriebene Prüflogik mit den benannten 5 Ebene ist nur dann nützlich, wenn daraus Entscheidungen entstehen. Wirksamkeitsmessung endet nicht bei der Datenerhebung, denn entscheidend ist, ob die erhobenen Daten so interpretiert werden, dass daraus konkrete Steuerungsimpulse abgeleitet werden können.

Ein wirksamkeitsorientierter Kennzahlenansatz muss deshalb drei Fragen beantworten:

Was zeigen die Daten?

Welche Entwicklungen sind bei Umsetzung, Transfer, Treibern, Arbeitserleben und Value-Faktoren sichtbar?

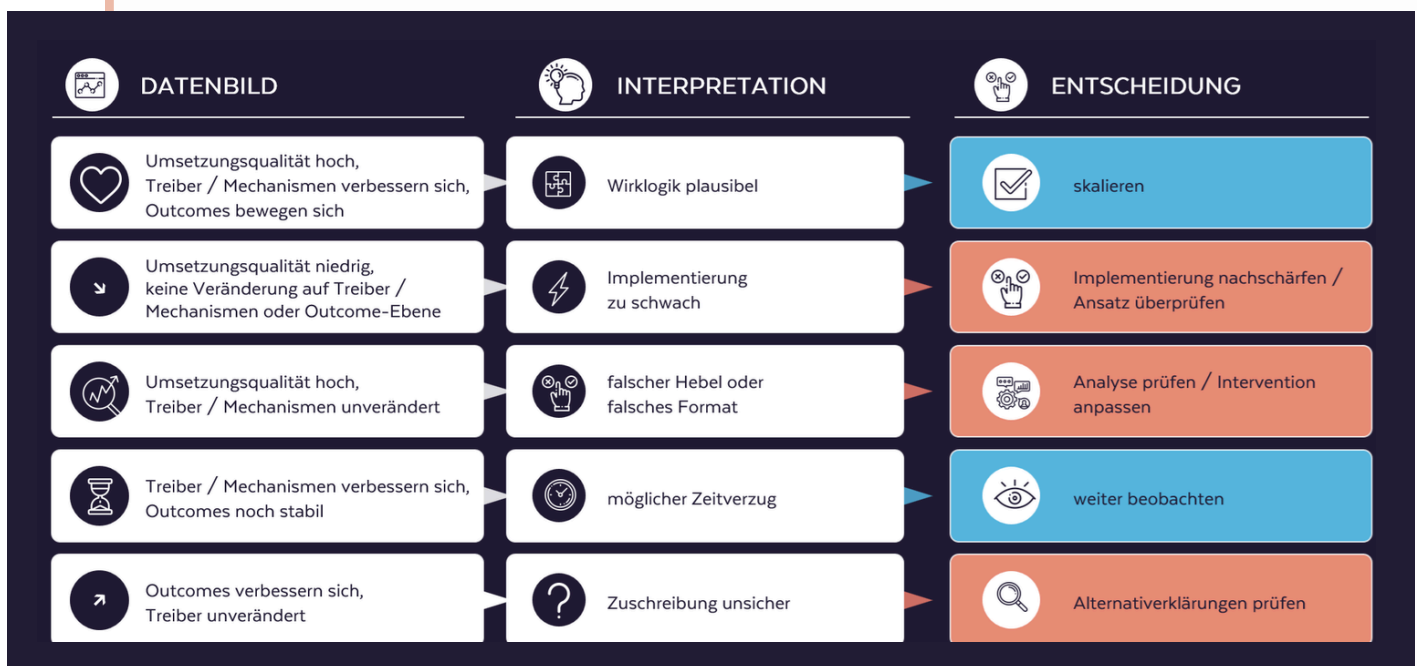
Wie sind diese Daten zu interpretieren?

Passen die Entwicklungen zur ursprünglichen Wirkungshypothese? Gibt es Widersprüche, Zeitverzug oder alternative Erklärungen?

Welche Entscheidung folgt daraus?

Soll die Maßnahme fortgeführt, angepasst, ausgeweitet, gestoppt oder zunächst weiter beobachtet werden?

Die folgende Steuerungslogik übersetzt typische Datenbilder direkt in Entscheidungen:



Der eigentliche Nutzen eines wirksamkeitsorientierten Kennzahlensystems liegt damit nicht im Bericht selbst, sondern in besseren Entscheidungen. Dadurch können Investitionen gezielter priorisiert, schwache Ansätze korrigiert, wirksame Ansätze ausgebaut und gegenüber HR und Geschäftsführung nachvollziehbarer begründet werden.

Wirksamkeitssteuerung bedeutet deshalb nicht, jede Veränderung endgültig zu beweisen, sondern Daten konsequent entlang einer Wirkungskette zu interpretieren und daraus fachlich begründete Entscheidungen abzuleiten.

Pragmatische Einstiege

Nach der Fünf-Ebenen-Prüflogik stellt sich die Frage, wie der Einstieg mit überschaubarem Aufwand gelingt. Der beste Startpunkt liegt grundsätzlich dort, wo bereits Handlungsdruck besteht, etwa bei erhöhten Fehlzeiten, steigender Fluktuation, Belastungsschwerpunkten aus der psychischen Gefährdungsbeurteilung oder offenen Fragen zur Wirkung laufender Maßnahmen.

Ausgangspunkt ist die Frage, welche Entscheidung in den nächsten Monaten besser getroffen werden soll. Daraus ergibt sich, welche Daten, Vergleiche und Routinen tatsächlich gebraucht werden. Häufig reichen vorhandene Informationen, eine klare Zielgruppe, eine einfache Wirkungskette und ein überschaubarer Pilot.

Die folgenden fünf Einstiege zeigen, wie Unternehmen pragmatisch in die Wirksamkeitssteuerung starten können.

Am Schmerzpunkt der Geschäftsführung ansetzen

Statt unternehmensweit zu starten, empfiehlt sich der Fokus auf eine Kennzahl in einem konkreten Bereich, die ohnehin bereits Aufmerksamkeit erhält – etwa erhöhte Fluktuation in der Nachtschicht, Krankenstand in der Logistik oder eine lange Wiederbesetzungsdauer im Vertrieb. Aus dieser Kennzahl wird rückwärts die Wirkungshypothese entwickelt, indem herausgearbeitet wird, welche Veränderungen im Arbeitserleben und bei den relevanten Treibern notwendig wären, um diese Zahl zu verändern. So entsteht ein Vorhaben, das von Beginn an anschlussfähig an die Unternehmensziele ist.

Vorhandene Daten heben, bevor neue erhoben werden

Bevor neue Daten erhoben werden, lohnt sich ein strukturierter Blick auf das, was bereits vorhanden ist. Austritts- und Bleibegespräche enthalten häufig wertvolle Hinweise zu Belastung, Führung und Bindung, gehen im Alltag jedoch oft in unstrukturierten Ablagen verloren. Fehlzeiten lassen sich in vielen Unternehmen bereits nach Bereich, Schicht oder Team betrachten, werden in der Steuerung aber häufig auf eine Gesamtquote reduziert. Auch die psychische Gefährdungsbeurteilung bietet einen guten Ausgangspunkt, weil sie Belastungsschwerpunkte vergleichsweise konkret benennt und nach der Maßnahmenableitung zu selten als wiederkehrende Steuerungsgrundlage genutzt wird.

Insbesondere die psychische Gefährdungsbeurteilung eignet sich als pragmatischer Einstieg in die Wirksamkeitssteuerung, weil sie gesetzlich verankert ist und in vielen Unternehmen bereits konkrete Hinweise auf Belastungsschwerpunkte liefert. Häufig endet der Prozess jedoch nach der Maßnahmenableitung, obwohl gerade die Nachverfolgung der identifizierten Belastungsfaktoren den eigentlichen Steuerungsnutzen schafft.

Dafür muss nicht die gesamte Erhebung wiederholt werden. Meist reicht ein kurzer Pulse Check nach drei bis sechs Monaten, der gezielt prüft, ob sich die bearbeiteten Belastungsfaktoren und Arbeitsbedingungen wie Arbeitsintensität, Rollenklarheit oder Erholung im Erleben der Beschäftigten verändert haben. So wird aus einer einmaligen Pflichtanalyse ein wiederkehrendes Steuerungsinstrument.

Der Vorteil liegt in der Verbindung von Arbeitsschutzlogik und Wirksamkeitssteuerung. Der Pulse Check kann die im Arbeitsschutzgesetz geforderte Wirksamkeitskontrolle der Maßnahmen unterstützen (§ 3 Abs. 1 ArbSchG) und zugleich zeigen, ob die priorisierten Belastungstreiber tatsächlich bearbeitet wurden.

Einen Leuchtturm statt eines Flächenrollouts

Der erste Wirksamkeitsnachweis muss nicht direkt unternehmensweit angelegt sein, sollte aber fachlich sauber und nachvollziehbar geführt werden. Oft reicht ein klar begrenzter Pilotfall, etwa ein Bereich, eine Maßnahme und eine durchgängige Wirkungskette, um intern zu zeigen, dass Wirksamkeitssteuerung praktisch funktioniert. Ein nachvollziehbarer Erfolg in einem konkreten Bereich schafft häufig mehr Rückendeckung und Folgebudget als ein umfassendes Konzept für viele Bereiche, das noch nicht erprobt wurde.

Den gestaffelten Rollout als Vergleichsgruppe nutzen

Wird ein Programm ohnehin nacheinander ausgerollt, kann daraus eine pragmatische Vergleichslogik entstehen. Zeigen sich Verbesserungen jeweils nach dem Start in den betreffenden Teams, spricht dies stärker für einen Zusammenhang mit der Maßnahme als eine reine Vorher-Nachher-Betrachtung.

Dafür sollten Rollout-Zeitpunkte, Zielgruppen, Messzeitpunkte und Kennzahlen sauber dokumentiert werden. Zusätzlich ist zu prüfen, ob Unterschiede zwischen den Teams oder parallele Entwicklungen die Ergebnisse ebenfalls erklären könnten.

Rhythmus vor Kennzahlen

Der unterschätzteste Hebel in der Wirksamkeitssteuerung ist oft kein zusätzliches Messinstrument, sondern ein fester Entscheidungsrythmus. Kennzahlen entfalten ihren Nutzen erst, wenn sie regelmäßig gemeinsam betrachtet, interpretiert und in konkrete Entscheidungen übersetzt werden.

Für den Einstieg reicht ein kompakter Quartalsreview mit drei Fragen.

1. Datenbild beschreiben

Was hat sich bei Umsetzung, Treibern, Arbeitserleben und Value-Faktoren verändert?

2. Datenbild interpretieren

Passt die Entwicklung zur angenommenen Wirkungshypothese? Gibt es Widersprüche, blinde Flecken oder alternative Erklärungen?

3. Entscheidung ableiten

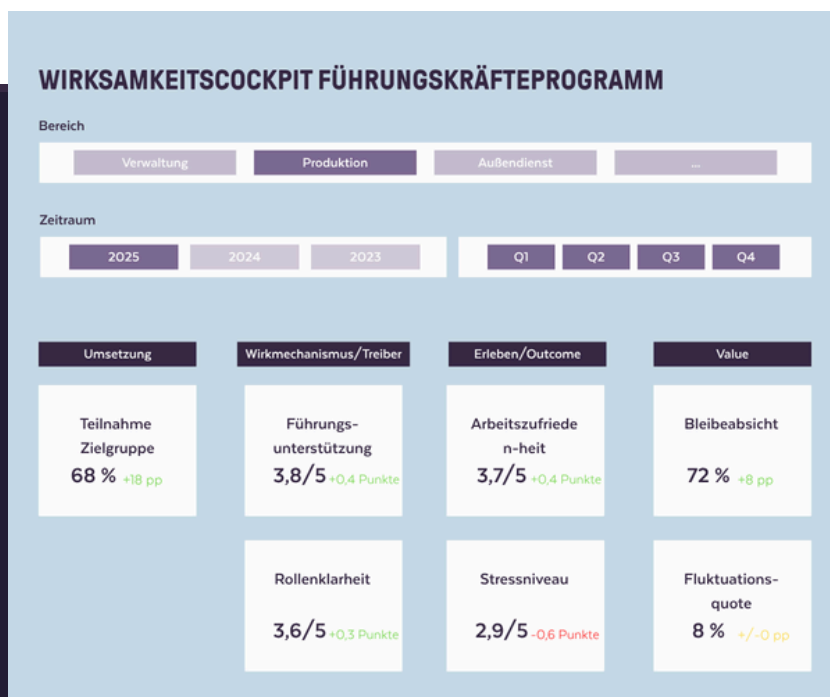
Soll die Maßnahme fortgeführt, nachgeschärft, ausgeweitet, gestoppt oder zunächst weiter beobachtet werden?

Gerade bei begrenzter Datenlage hilft ein regelmäßiger Steuerungstermin, einfache Kennzahlen wirksam zu nutzen, Entwicklungen gemeinsam einzuordnen und aus Messung konkrete Entscheidungsvorbereitung zu machen.

Ausblick - Vom Kennzahlenset zum Wirksamkeitscockpit

Der nächste Entwicklungsschritt besteht darin, die beschriebene Wirksamkeitslogik in ein einfach nutzbares Steuerungsinstrument zu übersetzen. Ein Wirksamkeitscockpit bündelt zentrale Datenpunkte und zeigt, wie sich Umsetzung, Treiber, Arbeitserleben und Value-Faktoren im Zeitverlauf entwickeln.

Der Nutzen liegt weniger in der einzelnen Kennzahl als im Gesamtbild. Am Beispiel des Führungskräfteprogramms (S. Abb) wird sichtbar, ob die relevante Zielgruppe erreicht wurde, ob sich Führungsunterstützung und Rollenklarheit verbessern, ob Veränderungen im Arbeitserleben erkennbar sind und ob sich Value-nahe Kennzahlen wie Bleibeabsicht oder Fluktuationsrisiko in die erwartete Richtung bewegen.



beispielhafte, schematische Darstellung

Entscheidend ist die gemeinsame Leserichtung. Links steht, was das Unternehmen unmittelbar gestalten und nachhalten kann. In der Mitte liegen Treiber und Arbeitserleben, auf denen Veränderungen häufig früher sichtbar werden. Rechts stehen die Value-Faktoren, die für HR, Geschäftsführung und Organisation besonders relevant sind. Durch diese Verbindung wird erkennbar, ob ein Programm die angenommene Wirkungskette plausibel in Bewegung setzt.

Damit ein Wirksamkeitscockpit mehr ist als ein Reporting-Instrument, braucht es eine einheitliche Filterlogik. Bereich, Zeitraum, Zielgruppe und Messlogik sollten über alle Ebenen hinweg konsistent abgebildet werden. Zugleich bleibt die Prüfung alternativer Erklärungen wichtig, weil auch ein positives Gesamtbild sorgfältig eingeordnet werden muss. Zu alternativen Erklärungen zählen beispielsweise saisonale Effekte, personelle Veränderungen, parallele Organisationsentwicklungen oder externe Einflüsse, die ebenfalls zu den beobachteten Veränderungen beigetragen haben können.

Seinen vollen Nutzen entfaltet ein Cockpit in einer festen Steuerungsroutine. In einem regelmäßigen Quartalsreview können HR, BGM, Führung und relevante Stakeholder gemeinsam prüfen, was sich verändert hat, ob die Entwicklung zur Wirkungshypothese passt und welche Entscheidung daraus folgt. Dadurch wird aus Messung konkrete Entscheidungsvorbereitung.

Fazit

Der Einstieg in eine Wirksamkeitslogik im Kontext Gesunder Arbeit kann für viele Unternehmen bereits mit wenigen grundlegenden Bausteine, wie klar definierten Zielen und Zielgruppen, einer nachvollziehbaren Wirkungshypothese und ausgewählten Kennzahlen, die regelmäßig betrachtet werden beginnen. Bereits damit lässt sich der wichtige Schritt von der reinen Aktivitätsmessung zu einer systematischeren Bewertung von Wirkung vollziehen.

Der Steuerungs- und Nachweis-Business-Case schafft dafür eine gemeinsame Logik. Maßnahmen und Lösungsarchitekturen werden entlang ihrer Wirkungskette, von den Voraussetzungen und der Umsetzungsqualität über Wirkmechanismen und Treiber bis zu Outcomes und organisationalen Value-Faktoren betrachtet. Dadurch wird sichtbar, wo eine Intervention die erwartete Entwicklung auslöst, an welcher Stelle es Probleme gibt und wo nachgesteuert werden sollte.

Mit jeder Anwendung wächst zugleich die Aussagefähigkeit. Ein Unternehmen kann zunächst mit vorhandenen Daten, wenigen Treibern und ausgewählten Outcomes beginnen und seine Wirksamkeitssteuerung schrittweise weiterentwickeln. Mit zunehmender Datenreife können längere Zeitreihen, Bereichsvergleiche und organisationale Value-Faktoren wie Absentismus, Präsentismus, Mitarbeiterbindung oder Produktivität stärker einbezogen werden. So entsteht aus einzelnen Messpunkten nach und nach eine belastbarere Grundlage für Steuerungs- und Investitionsentscheidungen.

Die Stärke dieses Ansatzes liegt in der Beitragslogik, bei der mehrere Datenpunkte entlang der zuvor formulierten Wirkungshypothese zusammengeführt und alternative Erklärungen systematisch eingeordnet werden. Je konsistenter sich die erwartete Entwicklung über die verschiedenen Ebenen zeigt, desto fundierter lässt sich beurteilen, welchen Beitrag eine Intervention geleistet hat und welche Konsequenzen daraus für ihre Fortführung, Anpassung oder Skalierung folgen.

Damit schließt sich der Kreis zum Entscheidungs-Business-Case. Die zunächst auf wissenschaftlicher Evidenz, Szenarien und Bandbreiten beruhenden Annahmen werden durch Erfahrungen und Unternehmensdaten überprüft und weiter geschärft. Aus der ursprünglichen Investitionsentscheidung entsteht so ein lernendes Steuerungssystem, das zukünftige Entscheidungen auf eine zunehmend belastbare Grundlage stellt.

Für HR und BGM liegt darin ein wichtiger Entwicklungsschritt, weil Gesunde Arbeit systematischer mit Arbeitsbedingungen, Arbeitserleben und organisationalen Ergebnissen verbunden und dadurch gegenüber Geschäftsführung und weiteren Stakeholdern fundierter gesteuert und begründet werden kann.

Der eigentliche Fortschritt liegt damit weniger in mehr Reporting als in besseren Entscheidungen darüber, wo investiert, nachgesteuert und weiterentwickelt werden sollte.

Quellen-/Literaturverzeichnis

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology*, 22(3), 309–328. <https://doi.org/10.1108/02683940710733115>

De Neve, J.-E., & Ward, G. (2023). Measuring workplace wellbeing (Working Paper 2303). Wellbeing Research Centre, University of Oxford. <https://wellbeing.hmc.ox.ac.uk/papers/wp-2303-measuring-workplace-wellbeing/>

De Neve, J.-E., & Ward, G. (2025). Why workplace wellbeing matters: The science behind employee happiness and organizational performance. Harvard Business Review Press.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542–575. <https://doi.org/10.1037/0033-2909.95.3.542>

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health*, 89(9), 1322–1327. <https://doi.org/10.2105/AJPH.89.9.1322>

Klingmann, T. (2026). Der Business Case Gesunde Arbeit. MOOVE-Whitepaper.

Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *Canadian Journal of Program Evaluation*, 16(1), 1–24. <https://doi.org/10.3138/cjpe.016.001>

Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O’Cathain, A., Tinati, T.,

Wight, D., & Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, 350, h1258. <https://doi.org/10.1136/bmj.h1258>

Nielsen, K., & Randall, R. (2013). Opening the black box: Presenting a model for evaluating organizational-level interventions. *European Journal of Work and Organizational Psychology*, 22(5), 601–617. <https://doi.org/10.1080/1359432X.2012.690556>

Niks, I. M. W., Veldhuis, G. A., van Zwieten, M. H. J., Sluijs, T., Wiezer, N. M., & Wortelboer, H. M. (2022). Individual workplace well-being captured into a literature- and stakeholders-based causal loop diagram. *International Journal of Environmental Research and Public Health*, 19(15), 8925. <https://doi.org/10.3390/ijerph19158925>

Rogers, P. (2014). Theory of Change (Methodological Briefs: Impact Evaluation No. 2). UNICEF Office of Research. https://www.betterevaluation.org/sites/default/files/Theory_of_Change_ENG.pdf